

Transmitter Optimization for Achieving Secrecy Capacity in Gaussian MIMO Wiretap Channels

Jiangyuan Li and Athina Petropulu

Department of Electrical and Computer Engineering

Drexel University, Philadelphia, PA 19104

Email: eejyli@yahoo.com.cn, athina@coe.drexel.edu

Abstract

We consider a Gaussian multiple-input multiple-output (MIMO) wiretap channel model, where there exists a transmitter, a legitimate receiver and an eavesdropper, each node equipped with multiple antennas. We study the problem of finding the optimal input covariance matrix that achieves secrecy capacity subject to a power constraint, which leads to a non-convex optimization problem that is in general difficult to solve. Existing results for this problem address the case in which the transmitter and the legitimate receiver have two antennas each and the eavesdropper has one antenna. For the general cases, it has been shown that the optimal input covariance matrix has low rank when the difference between the Grams of the eavesdropper and the legitimate receiver channel matrices is indefinite or semi-definite, while it may have low rank or full rank when the difference is positive definite. In this paper, the aforementioned non-convex optimization problem is investigated. In particular, for the multiple-input single-output (MISO) wiretap channel, the optimal input covariance matrix is obtained in closed form. For general cases, we derive the necessary conditions for the optimal input covariance matrix consisting of a set of equations. For the case in which the transmitter has two antennas, the derived necessary conditions can result in a closed form solution; For the case in which the difference between the Grams is indefinite and has all negative eigenvalues except one positive eigenvalue, the optimal input covariance matrix has rank one and can be obtained in closed form; For other cases, the solution is proved to be a fixed point of a mapping from a convex set to itself and an iterative procedure is provided to search for it. Numerical results are presented to illustrate the proposed theoretical findings.

Index Terms

Secrecy capacity, Gaussian MIMO wiretap channel, transmitter optimization, physical layer based security.

I. INTRODUCTION

Wireless physical (PHY) layer based security from an information-theoretic point of view has received considerable attention recently, e.g., [1]-[5], and the comprehensive overview in [6]. Wireless PHY layer based security approaches exploit the physical characteristics of the wireless channel to enhance the security of communication systems. The wiretap channel, first introduced and studied by Wyner [7], is the most basic physical layer model that captures the problem of communication security. Wyner showed that when an eavesdropper's channel is a degraded version of the main channel, the source and destination can achieve a positive perfect information rate (secrecy rate). The maximal rate of secrecy rate from the source to the destination is defined as the *secrecy capacity* and for the degraded wiretap channel is given as the largest between zero and the difference between the capacity at the legitimate receiver and the capacity at the eavesdropper. The Gaussian wiretap channel, in which the outputs at the legitimate receiver and at the eavesdropper are corrupted by additive white Gaussian noise (AWGN), was studied in [8]. Along the same line, the Gaussian MIMO wiretap channel was investigated and the secrecy capacity of the MIMO wiretap channel was established in terms of an optimization problem over all possible input covariance matrices [9], [10]. In [9], [10], the Gaussian MIMO wiretap channel model was given as $\mathbf{y}_i = \mathbf{H}_i \mathbf{x} + \mathbf{n}_i$, $i = 1, 2$ where \mathbf{n}_i is AWGN with zero mean and covariance $\sigma^2 \mathbf{I}$, and the power constraint $\text{Tr}(\mathbf{R}_x) \leq P$ was used, where \mathbf{R}_x is the input covariance matrix. An alternative expression of secrecy capacity was derived in [11], [12] for another Gaussian MIMO wiretap channel model, i.e., $\mathbf{y}_i = \mathbf{x} + \mathbf{v}_i$, $i = 1, 2$ where \mathbf{v}_i is additive Gaussian noise (AGN) with zero mean and invertible covariance \mathbf{W}_i . Further, in [12], the power covariance constraint $\mathbf{R}_x \preceq \mathbf{S}$ was used where \mathbf{S} is a given matrix and $\mathbf{R}_x \preceq \mathbf{S}$ denotes that $\mathbf{S} - \mathbf{R}_x$ is positive semi-definite, which allowed for the secrecy capacity to be obtained in closed form rather than as a solution to an optimization problem.

For the former Gaussian MIMO wiretap channel model, i.e., $\mathbf{y}_i = \mathbf{H}_i \mathbf{x} + \mathbf{n}_i$ and the power constraint $\text{Tr}(\mathbf{R}_x) \leq P$, finding the optimal input covariance matrix that achieves secrecy capacity leads to a non-convex optimization problem. This problem is in general difficult to solve. The solution of a special case in which the transmitter and the legitimate receiver each has two antennas and the eavesdropper has one antenna was given in [13]. In [14], it was pointed out that the optimal input covariance matrix has low rank when the difference between the Grams of the eavesdropper and the legitimate receiver channel matrices is indefinite or semi-definite, based on the assumption that the Grams both have full rank.

In this paper, we investigate the aforementioned non-convex optimization problem. In particular, for the multiple-input single-output (MISO) wiretap channel, we obtain the optimal input covariance matrix in closed form. For general MIMO case, we derive the necessary conditions for the optimal solution consisting of a set of equations. Those conditions result in a closed form solution for $n_T = 2$. For the more general case in which the difference of the Grams of the eavesdropper and the legitimate receiver channel matrices is indefinite and has all negative eigenvalues except one positive eigenvalue, we prove that the optimal input covariance matrix has rank one and can be obtained in closed form. Otherwise, we prove that the solution is a fixed point of a mapping from a convex set to itself and provide an iterative procedure to search for it.

The remainder of this paper is organized as follows. The mathematical model is introduced in §II. In §III the optimal input covariance matrix is obtained in closed form for Gaussian MISO wiretap channel. In §IV we derive the necessary conditions for the optimal solution consisting of a set of equations for the Gaussian MIMO wiretap channel. In §V, we obtain a closed form solution for the case in which the transmitter has two antennas. In §VI, for the case in which the difference of the Grams is indefinite and has all negative eigenvalues except one positive eigenvalue, we prove that the optimal input covariance matrix has rank one and can be obtained in closed form. Numerical results in §VIII illustrate the proposed algorithm. Finally, §IX gives a brief conclusion.

A. Notation

Upper case and lower case bold symbols denote matrices and vectors, respectively. Superscripts $*$, T and \dagger denote respectively conjugate, transposition and conjugate transposition. $\det(\mathbf{A})$ and $\text{Tr}(\mathbf{A})$ denote the determinant and trace of matrix \mathbf{A} , respectively. $\lambda_{\max}(\mathbf{A})$ denotes the largest eigenvalue of \mathbf{A} . $\mathbf{A} \succeq 0$ and $\mathbf{A} \succ 0$ mean that \mathbf{A} is a Hermitian positive semi-definite and positive definite matrix, respectively. $\mathbf{A} \succeq \mathbf{B}$ denotes that $\mathbf{A} - \mathbf{B}$ is a positive semi-definite matrix. $\mathbf{A} \not\succ 0$ denotes that matrix \mathbf{A} is not positive definite. $\text{rank}(\mathbf{A})$ denotes the rank of matrix \mathbf{A} . $\text{diag}(\mathbf{v})$ denotes a diagonal matrix with diagonal entries consisting of the elements of \mathbf{v} . $\|\mathbf{a}\|$ denotes Euclidean norm of vector \mathbf{a} , while $\|\mathbf{A}\|$ denotes Frobenius norm of matrix \mathbf{A} . \mathbf{I}_n denotes the identity matrix of order n (the subscript is dropped when the dimension is obvious). $a \rightarrow b$ means that a goes to b . Given a matrix \mathbf{A} , the matrix $\mathbf{A}^\dagger \mathbf{A}$ is called a Gram.

II. SYSTEM MODEL AND FORMULATIONS

Consider a MIMO wiretap channel where the transmitter is equipped with n_T antennas, while the legitimate receiver and an eavesdropper have n_R and n_E antennas, respectively. The received signals at the legitimate receiver and the eavesdropper are respectively given by

$$\mathbf{y}_R = \mathbf{H}_R \mathbf{x} + \mathbf{n}_R, \mathbf{y}_E = \mathbf{H}_E \mathbf{x} + \mathbf{n}_E \quad (1)$$

where \mathbf{H}_R ($n_R \times n_T$), \mathbf{H}_E ($n_E \times n_T$) are respectively channel matrices between the transmitter and legitimate receiver, and between the transmitter and eavesdropper; \mathbf{x} is the $n_T \times 1$ transmitted signal vector with zero mean and $n_T \times n_T$ covariance matrix $\mathbf{R}_x \succeq 0$; \mathbf{n}_R and \mathbf{n}_E are circular Gaussian noise vectors with zero mean and covariance matrices $\sigma^2 \mathbf{I}_{n_R}$ and $\sigma^2 \mathbf{I}_{n_E}$, respectively. We assume the power constraint is P , namely, $\text{Tr}(\mathbf{R}_x) = P$. It is easy to verify that the problem under $\text{Tr}(\mathbf{R}_x) \leq P$ is equivalent to that under $\text{Tr}(\mathbf{R}_x) = P$. We consider the scenario in which the transmitter has perfect short-term channel state information (CSI).

The secrecy capacity is defined as [10]

$$C_s \triangleq \max_{\mathbf{R}_x \succeq 0, \text{Tr}(\mathbf{R}_x)=P} C_s(\mathbf{R}_x) \quad (2)$$

where

$$C_s(\mathbf{R}_x) = \log \det(\mathbf{I}_{n_R} + \mathbf{H}_R \mathbf{R}_x \mathbf{H}_R^\dagger / \sigma^2) - \log \det(\mathbf{I}_{n_E} + \mathbf{H}_E \mathbf{R}_x \mathbf{H}_E^\dagger / \sigma^2) \quad (3)$$

is the secrecy rate.

The transmitter optimization problem is to determine \mathbf{R}_x that maximizes the secrecy rate, i.e., achieves secrecy capacity. The optimization makes sense when the secrecy capacity is positive. We assume $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E \neq 0$ since otherwise $C_s(\mathbf{R}_x) \equiv 0$. Whether $C_s > 0$ depends on the difference between the Grams of the legitimate receiver and eavesdropper channel matrices, i.e., $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E$. The following lemma provides the conditions to maintain $C_s > 0$.

Lemma 1: *The sufficient and necessary condition for $C_s > 0$ is: $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E$ is positive semi-definite or indefinite.*

Please see Appendix A for details.

Let us assume that the channel matrices \mathbf{H}_R and \mathbf{H}_E have been normalized so that $\text{Tr}(\mathbf{H}_R^\dagger \mathbf{H}_R) = n_T$. We denote the signal-to-noise ratio (SNR) $\rho \triangleq P/\sigma^2$ and let $\mathbf{R}_x = P\mathbf{Q}$. The constraints in (2) now become $\mathbf{Q} \succeq 0$, $\text{Tr}(\mathbf{Q}) = 1$. The secrecy rate maximization problem can be written as

$$\max_{\mathbf{Q} \succeq 0, \text{Tr}(\mathbf{Q})=1} C_s(\mathbf{Q}) = \log \det(\mathbf{I}_{n_R} + \rho \mathbf{H}_R \mathbf{Q} \mathbf{H}_R^\dagger) - \log \det(\mathbf{I}_{n_E} + \rho \mathbf{H}_E \mathbf{Q} \mathbf{H}_E^\dagger). \quad (4)$$

According to Lemma 1, we assume $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E$ is positive semi-definite or indefinite (except the MISO channel). In [14], the authors assumed that $\mathbf{H}_R^\dagger \mathbf{H}_R$ and $\mathbf{H}_E^\dagger \mathbf{H}_E$ are both positive definite, and hence they are both invertible. Here, we do not make that assumption. In fact, when $n_T > n_R$ (and/or $n_T > n_E$), $\mathbf{H}_R^\dagger \mathbf{H}_R$ (and/or $\mathbf{H}_E^\dagger \mathbf{H}_E$) always have low rank and hence are not invertible.

We also denote the feasible set of (4) as

$$\Omega = \{\mathbf{Q} | \mathbf{Q} \succeq 0, \text{Tr}(\mathbf{Q}) = 1\} \quad (5)$$

which is a convex set.

III. CLOSED FORM SECRECY CAPACITY OF MISO WIRETAP CHANNEL

We first provide a lemma that will be used here and in the proof of Theorem 6 later.

Lemma 2: *Let \mathbf{r} and \mathbf{s} be two known non-zero vectors, and $\mathbf{r}\mathbf{r}^\dagger - \mathbf{s}\mathbf{s}^\dagger \neq 0$.*

- (i) *If $\mathbf{r} = \xi \mathbf{s}$ for a certain scalar ξ , $\mathbf{r}\mathbf{r}^\dagger - \mathbf{s}\mathbf{s}^\dagger$ has only one nonzero eigenvalue equal to $(|\xi|^2 - 1)\|\mathbf{s}\|^2$ with the associated eigenvector $\mathbf{s}/\|\mathbf{s}\|$;*
- (ii) *If $\mathbf{r}^\dagger \mathbf{s} = 0$, $\mathbf{r}\mathbf{r}^\dagger - \mathbf{s}\mathbf{s}^\dagger$ has only two nonzero eigenvalues, i.e., $\eta_1 = \|\mathbf{r}\|^2$, $\eta_2 = -\|\mathbf{s}\|^2$ with associated eigenvectors $\mathbf{r}/\|\mathbf{r}\|$, $\mathbf{s}/\|\mathbf{s}\|$, respectively.*
- (iii) *If neither $\mathbf{r} = \xi \mathbf{s}$ nor $\mathbf{r}^\dagger \mathbf{s} = 0$, $\mathbf{r}\mathbf{r}^\dagger - \mathbf{s}\mathbf{s}^\dagger$ has only two nonzero eigenvalues, i.e., $\eta_1 = \|\mathbf{r}\|^2 - |c_2||\mathbf{r}^\dagger \mathbf{s}| > 0$, $\eta_2 = \|\mathbf{r}\|^2 - |c_4||\mathbf{r}^\dagger \mathbf{s}| < 0$ with the associated eigenvectors $\mathbf{e}_1 = c_1^{-1/2}(\mathbf{r} + |c_2|e^{i(\pi-\varphi)}\mathbf{s})$, $\mathbf{e}_2 = c_3^{-1/2}(\mathbf{r} + |c_4|e^{i(\pi-\varphi)}\mathbf{s})$, respectively, where φ is the argument of $\mathbf{r}^\dagger \mathbf{s}$, $i = \sqrt{-1}$, $c_1 = \|\mathbf{r}\|^2 + |c_2|^2\|\mathbf{s}\|^2 - 2|c_2||\mathbf{r}^\dagger \mathbf{s}|$, $|c_2| = (\|\mathbf{r}\|^2 + \|\mathbf{s}\|^2 - \sqrt{(\|\mathbf{r}\|^2 + \|\mathbf{s}\|^2)^2 - 4|\mathbf{r}^\dagger \mathbf{s}|^2})/(2|\mathbf{r}^\dagger \mathbf{s}|)$, $c_3 = \|\mathbf{r}\|^2 + |c_4|^2\|\mathbf{s}\|^2 - 2|c_4||\mathbf{r}^\dagger \mathbf{s}|$, $|c_4| = (\|\mathbf{r}\|^2 + \|\mathbf{s}\|^2 + \sqrt{(\|\mathbf{r}\|^2 + \|\mathbf{s}\|^2)^2 - 4|\mathbf{r}^\dagger \mathbf{s}|^2})/(2|\mathbf{r}^\dagger \mathbf{s}|)$.*

The proof is simple, therefore, omitted for the sake of brevity. But we outline the proof here. For the case (i), (ii), the proof is obvious. For the case (iii), first, we can show $\mathbf{r}\mathbf{r}^\dagger - \mathbf{s}\mathbf{s}^\dagger$ has rank two, thus it has only two nonzero eigenvalues. Second, we assume the eigenvector has the form of a linear combination of \mathbf{r} and \mathbf{s} , and then show that this is indeed the case.

Before discussing the general MIMO wiretap channel, we analyze a special case, i.e., the MISO wiretap channel in which the legitimate receiver and eavesdropper both have a single antenna, i.e., $n_R = n_E = 1$. Denote the channel vectors as \mathbf{h}_R and \mathbf{h}_E . We give the following theorem.

Theorem 1: *The closed form expression for secrecy capacity of MISO wiretap channel is given by*

$$C_s = \log \frac{b + \sqrt{b^2 - 4ac}}{2a} \quad (6)$$

where $a = 1 + \rho\|\mathbf{h}_E\|^2$, $b = 2 + \rho\|\mathbf{h}_R\|^2 + \rho\|\mathbf{h}_E\|^2 + \rho^2(\|\mathbf{h}_R\|^2\|\mathbf{h}_E\|^2 - |\mathbf{h}_R^\dagger \mathbf{h}_E|^2)$ and $c = 1 + \rho\|\mathbf{h}_R\|^2$.

Proof: The secrecy rate maximization problem can be written as

$$\max_{\mathbf{Q} \succeq 0, \text{Tr}(\mathbf{Q})=1} C_s(\mathbf{Q}) = \log \frac{1 + \rho \mathbf{h}_R^\dagger \mathbf{Q} \mathbf{h}_R}{1 + \rho \mathbf{h}_E^\dagger \mathbf{Q} \mathbf{h}_E} \quad (7)$$

which is a fractional program [20] associated with the following parametric problem

$$F(\alpha) = \max_{\mathbf{Q} \succeq 0, \text{Tr}(\mathbf{Q})=1} \left[1 + \rho \mathbf{h}_R^\dagger \mathbf{Q} \mathbf{h}_R - \alpha(1 + \rho \mathbf{h}_E^\dagger \mathbf{Q} \mathbf{h}_E) \right] \quad (8)$$

where $\alpha > 0$. Let α° be the unique root of $F(\alpha) = 0$. According to [20], the optimal \mathbf{Q} corresponding to $F(\alpha^\circ)$ also optimizes (7). Based on the fact that $\mathbf{h}^\dagger \mathbf{Q} \mathbf{h} = \text{Tr}(\mathbf{Q} \mathbf{h} \mathbf{h}^\dagger)$ for any vector \mathbf{h} , we rewrite the optimization problem (8) as

$$F(\alpha) = \max_{\mathbf{Q} \succeq 0, \text{Tr}(\mathbf{Q})=1} \left[1 - \alpha + \rho \text{Tr}\{\mathbf{Q}(\mathbf{h}_R \mathbf{h}_R^\dagger - \alpha \mathbf{h}_E \mathbf{h}_E^\dagger)\} \right]. \quad (9)$$

By eigen-decomposition $\mathbf{h}_R \mathbf{h}_R^\dagger - \alpha \mathbf{h}_E \mathbf{h}_E^\dagger = \mathbf{U}_\alpha \mathbf{D}_\alpha \mathbf{U}_\alpha^\dagger$ and letting $\mathbf{Q}_\alpha = \mathbf{U}_\alpha^\dagger \mathbf{Q} \mathbf{U}_\alpha$, we obtain $\mathbf{Q}_\alpha \succeq 0$, $\text{Tr}(\mathbf{Q}_\alpha) = 1$, $\mathbf{Q} = \mathbf{U}_\alpha \mathbf{Q}_\alpha \mathbf{U}_\alpha^\dagger$. It holds

$$\text{Tr}\{\mathbf{Q}(\mathbf{h}_R \mathbf{h}_R^\dagger - \alpha \mathbf{h}_E \mathbf{h}_E^\dagger)\} = \text{Tr}\{\mathbf{Q}_\alpha \mathbf{D}_\alpha\} = \text{Tr}\{\text{diag}(\mathbf{Q}_\alpha) \mathbf{D}_\alpha\} \leq \lambda_{\max}(\mathbf{h}_R \mathbf{h}_R^\dagger - \alpha \mathbf{h}_E \mathbf{h}_E^\dagger). \quad (10)$$

Equation (10) holds with equality if \mathbf{Q}_α is diagonal and has a unique nonzero entry (equal to one) corresponding to position of the largest entry in \mathbf{D}_α . In other words, \mathbf{Q} and $\mathbf{h}_R \mathbf{h}_R^\dagger - \alpha \mathbf{h}_E \mathbf{h}_E^\dagger$ have the same eigenvectors, and \mathbf{Q} has rank one. Thus, it holds $\mathbf{Q} = \mathbf{u}_{\alpha, \max} \mathbf{u}_{\alpha, \max}^\dagger$ where $\mathbf{u}_{\alpha, \max}$ is the eigenvector associated with the largest eigenvalue of $\mathbf{h}_R \mathbf{h}_R^\dagger - \alpha \mathbf{h}_E \mathbf{h}_E^\dagger$. The largest eigenvalue and the associated eigenvector of $\mathbf{h}_R \mathbf{h}_R^\dagger - \alpha \mathbf{h}_E \mathbf{h}_E^\dagger$ can be expressed in closed form based on Lemma 2. In our problem, $\mathbf{r} = \mathbf{h}_R$, $\mathbf{s} = \sqrt{\alpha} \mathbf{h}_E$. By using Lemma 2, we now can obtain

$$F(\alpha) = 1 + \frac{\rho \|\mathbf{h}_R\|^2}{2} - \left(1 + \frac{\rho \|\mathbf{h}_E\|^2}{2} \right) \alpha + \frac{\rho}{2} \sqrt{(\|\mathbf{h}_R\|^2 + \alpha \|\mathbf{h}_E\|^2)^2 - 4\alpha |\mathbf{h}_R^\dagger \mathbf{h}_E|^2}. \quad (11)$$

$F(\alpha) = 0$ has a unique root given in closed form:

$$\alpha^\circ = \frac{b + \sqrt{b^2 - 4ac}}{2a}. \quad (12)$$

The optimal \mathbf{Q} is given by $\mathbf{Q}^\circ = \mathbf{e}_1 \mathbf{e}_1^\dagger$ where \mathbf{e}_1 is defined in Lemma 2 where $\mathbf{r} = \mathbf{h}_R$, $\mathbf{s} = \sqrt{\alpha^\circ} \mathbf{h}_E$. The secrecy capacity is given by $C_s = \log \alpha^\circ$. ■

Based on Theorem 1, if $\mathbf{h}_R = \xi \mathbf{h}_E$ and $|\xi| < 1$, then $b = a + c$, $a - c > 0$ and further $C_s = 0$. This is consistent with the fact that when the legitimate receiver channel is a degraded version of the eavesdropper channel the secrecy capacity is zero. If $\mathbf{h}_R = \xi \mathbf{h}_E$ and $|\xi| > 1$, then $b = a + c$, $a - c < 0$ and $C_s = \log((1 + \rho |\xi|^2 \|\mathbf{h}_E\|^2)/(1 + \rho \|\mathbf{h}_E\|^2)) > 0$. This is consistent with the fact that when the

eavesdropper channel is a degraded version of the legitimate receiver channel the secrecy capacity is positive. If $\mathbf{h}_R \neq \xi \mathbf{h}_E$, then $b > a + c$ and it always holds that $C_s > 0$. Thus, if $\mathbf{h}_R \neq \xi \mathbf{h}_E$, the MISO wiretap channel always has positive secrecy capacity independent of the channel.

To gain more insight into the secrecy capacity, we consider the rate at which the secrecy capacity scales with $\log \rho$ as in [15]. If $\mathbf{h}_R \neq \xi \mathbf{h}_E$, then under high SNR, it follows from (6) that

$$C_s(\rho) = \log \rho + \log(\|\mathbf{h}_R\|^2 - |\mathbf{h}_R^\dagger \mathbf{h}_E|^2 / \|\mathbf{h}_E\|^2 + O(1/\rho)) \quad (13)$$

where $O(\cdot)$ is the big-O notation. The secrecy degree of freedom (*s.d.o.f.*) (also see [15]) of the MISO wiretap channel is given by

$$s.d.o.f. \triangleq \lim_{\rho \rightarrow \infty} \frac{C_s(\rho)}{\log \rho} = 1. \quad (14)$$

IV. CONDITIONS FOR OPTIMAL INPUT COVARIANCE MATRIX OF MIMO WIRETAP CHANNEL

In this section, we analyze a general MIMO wiretap channel. First, we obtain the necessary conditions for the optimal \mathbf{Q} by using Karush-Kuhn-Tucker (KKT) conditions. Let us construct the cost function

$$L(\mathbf{Q}, \theta, \Psi) = C_s(\mathbf{Q}) - \theta(\text{Tr}(\mathbf{Q}) - 1) + \text{Tr}(\Psi \mathbf{Q}) \quad (15)$$

where θ is the Lagrange multiplier associated with the constraint $\text{Tr}(\mathbf{Q}) = 1$, Ψ is the Lagrange multiplier associated with the constraint $\mathbf{Q} \succeq 0$. The KKT conditions enable us to write [21]

$$\Theta - \theta \mathbf{I}_{n_T} + \Psi = 0, \quad (16)$$

$$\Psi \succeq 0, \text{Tr}(\Psi \mathbf{Q}) = 0, \mathbf{Q} \succeq 0, \text{Tr}(\mathbf{Q}) = 1, \quad (17)$$

where

$$\Theta = \rho \mathbf{H}_R^\dagger (\mathbf{I}_{n_R} + \rho \mathbf{H}_R \mathbf{Q} \mathbf{H}_R^\dagger)^{-1} \mathbf{H}_R - \rho \mathbf{H}_E^\dagger (\mathbf{I}_{n_E} + \rho \mathbf{H}_E \mathbf{Q} \mathbf{H}_E^\dagger)^{-1} \mathbf{H}_E. \quad (18)$$

Here we use the facts: $\frac{\partial}{\partial \mathbf{Q}} \text{Tr}(\Psi \mathbf{Q}) = \Psi^T$ and

$$\frac{\partial \log \det(\mathbf{I}_{n_R} + \rho \mathbf{H}_R \mathbf{Q} \mathbf{H}_R^\dagger)}{\partial \mathbf{Q}} = [\rho \mathbf{H}_R^\dagger (\mathbf{I}_{n_R} + \rho \mathbf{H}_R \mathbf{Q} \mathbf{H}_R^\dagger)^{-1} \mathbf{H}_R]^T. \quad (19)$$

For future use, we also rewrite (18) as

$$\Theta = \mathbf{S}_R (\mathbf{I}_{n_T} + \mathbf{Q} \mathbf{S}_R)^{-1} - \mathbf{S}_E (\mathbf{I}_{n_T} + \mathbf{Q} \mathbf{S}_E)^{-1} \quad (20)$$

where $\mathbf{S}_R = \rho \mathbf{H}_R^\dagger \mathbf{H}_R$, $\mathbf{S}_E = \rho \mathbf{H}_E^\dagger \mathbf{H}_E$, which follows from the fact: $\mathbf{H}(\mathbf{I} + \rho \mathbf{Q} \mathbf{H}^\dagger \mathbf{H}) = (\mathbf{I} + \rho \mathbf{H} \mathbf{Q} \mathbf{H}^\dagger) \mathbf{H}$ for any matrix \mathbf{H} and hence $(\mathbf{I} + \rho \mathbf{H} \mathbf{Q} \mathbf{H}^\dagger)^{-1} \mathbf{H} = \mathbf{H}(\mathbf{I} + \rho \mathbf{Q} \mathbf{H}^\dagger \mathbf{H})^{-1}$, $\mathbf{H}^\dagger (\mathbf{I} + \rho \mathbf{H} \mathbf{Q} \mathbf{H}^\dagger)^{-1} \mathbf{H} = \mathbf{H}^\dagger \mathbf{H} (\mathbf{I} + \rho \mathbf{Q} \mathbf{H}^\dagger \mathbf{H})^{-1}$.

In this paper, Θ is an important variable for the optimal input covariance problem. It has the following property which will be used later.

Property 1: *For any $\mathbf{Q} \succeq 0$, $\lambda_{\max}(\Theta) > 0$; For any $\mathbf{Q} \succeq 0$, $\text{Tr}(\mathbf{Q}\Theta) \leq \text{Tr}(\mathbf{Q})\lambda_{\max}(\Theta)$, and in particular, for any $\mathbf{Q} \in \Omega$, $\text{Tr}(\mathbf{Q}\Theta) \leq \lambda_{\max}(\Theta)$.*

The proof is given in Appendix B.

From the KKT conditions (16) and (17), we obtain the equivalent (but without containing the Lagrange multipliers) conditions for optimal \mathbf{Q} consisting of a set of equations given in the following theorem.

Theorem 2: *The optimal $\mathbf{Q} \succeq 0$ satisfies*

$$\mathbf{Q}\Theta = \text{Tr}(\mathbf{Q}\Theta)\mathbf{Q} \quad (21)$$

$$\lambda_{\max}(\Theta) = \text{Tr}(\mathbf{Q}\Theta). \quad (22)$$

Please see Appendix C for details.

Equations (21) and (22) provide two elementary conditions that characterize the optimal \mathbf{Q} . At this point we do not have a proof that any \mathbf{Q} satisfying the conditions of Theorem 2 is the optimal input covariance. However, for some special cases, e.g., the MISO wiretap channel analyzed in §III, this is true. In particular, for this case we provide the following theorem.

Theorem 3: *For MISO wiretap channel, any \mathbf{Q} satisfying the conditions of Theorem 2 is the optimal input covariance.*

The proof is given in Appendix D.

Now we proceed. From Property 1 and (22), we know that the optimal \mathbf{Q} satisfies

$$\text{Tr}(\mathbf{Q}\Theta) > 0. \quad (23)$$

Based on (21) and (23), and by taking trace operation over both side of (21), it can be easily seen that $\text{Tr}(\mathbf{Q}) = 1$. That is to say, equations (21) and (22) imply $\text{Tr}(\mathbf{Q}\Theta) > 0$ and $\text{Tr}(\mathbf{Q}) = 1$.

The condition (21) reveals that the optimal \mathbf{Q} satisfies that \mathbf{Q} and Θ commute and have the same eigenvectors [22, p.239]. The condition (22) means that the eigenvalues of Θ corresponding to the positive eigenvalues of \mathbf{Q} are all equal to $\text{Tr}(\mathbf{Q}\Theta)$, while the remaining eigenvalues of Θ (i.e., corresponds to the zero eigenvalues of \mathbf{Q}) are all less than or equal to $\text{Tr}(\mathbf{Q}\Theta)$. Obviously, if the optimal \mathbf{Q} has full rank, then $\Theta = \theta \mathbf{I}_{n_T}$ for a certain $\theta > 0$.

It can be shown that based on the conditions of Theorem 2, the optimal \mathbf{Q} has the following properties.

Property 2: *The optimal \mathbf{Q} satisfies:*

- (i) $\text{rank}(\Theta) = \text{rank}(\mathbf{S}_R - \mathbf{S}_E) \geq \text{rank}(\mathbf{Q})$;
- (ii) $\mathbf{Q}(\mathbf{S}_R - \mathbf{S}_E)\mathbf{Q} \succeq 0$;
- (iii) $\mathbf{Q} + \mathbf{Q}\mathbf{S}_E\mathbf{Q}$ and $\mathbf{Q} + \mathbf{Q}\mathbf{S}_R\mathbf{Q}$ commute and have the same eigenvectors.

For readability, we put the proof of Property 2 in Appendix E. A direct result of Property 2 is the following:

Property 3: *when $\mathbf{S}_R - \mathbf{S}_E \not\succeq 0$, the optimal \mathbf{Q} has low rank.*

The proof is simple. When $\mathbf{S}_R - \mathbf{S}_E$ is indefinite, if the optimal \mathbf{Q} has full rank, then Property 2 (ii) leads to $\mathbf{S}_R - \mathbf{S}_E \succeq 0$, which violates that $\mathbf{S}_R - \mathbf{S}_E$ is indefinite. When $\mathbf{S}_R - \mathbf{S}_E \succeq 0$ but $\mathbf{S}_R - \mathbf{S}_E \not\succeq 0$, it follows from Property 2 (i) that the optimal \mathbf{Q} has low rank. This result was also pointed out in [14]. When $\mathbf{S}_R - \mathbf{S}_E \succ 0$, the optimal \mathbf{Q} may have low rank or full rank.

Before ending this section, we point out that we can combine the elementary conditions (21) and (22) into a single equation. When Θ and \mathbf{Q} commute and have the same eigenvectors, $\Theta + \gamma\mathbf{I}_{n_T}$ and \mathbf{Q} commute and have the same eigenvectors for any real number γ , and vice versa. We can find a certain γ such that $\Theta + \gamma\mathbf{I}_{n_T} \succ 0$ for any $\mathbf{Q} \succeq 0$. Based on Θ in (20), we have

$$\lambda_{\min}(\Theta) > -\lambda_{\max}(\mathbf{S}_E). \quad (24)$$

When $\gamma \geq \lambda_{\max}(\mathbf{S}_E)$, it always holds that $\Theta + \gamma\mathbf{I}_{n_T} \succ 0$. Let $\mathbf{K} = \Theta + \gamma\mathbf{I}_{n_T}$ and hence $\mathbf{K} \succ 0$, $\text{Tr}(\mathbf{Q}\mathbf{K}) > 0$ for any $\mathbf{Q} \succeq 0$ but $\mathbf{Q} \neq 0$. Equations (21) and (22) are equivalent to

$$\mathbf{Q}\mathbf{K} = \text{Tr}(\mathbf{Q}\mathbf{K})\mathbf{Q}, \quad (25)$$

$$\lambda_{\max}(\mathbf{K}) = \text{Tr}(\mathbf{Q}\mathbf{K}). \quad (26)$$

We can combine the above two equations to a single one as follows.

$$\mathbf{Q}\mathbf{K} = \frac{1}{2} \left(\text{Tr}(\mathbf{Q}) + \frac{1}{\text{Tr}(\mathbf{Q})} \right) \lambda_{\max}(\mathbf{K})\mathbf{Q} \quad (27)$$

or equivalently,

$$\mathbf{Q}\tilde{\mathbf{K}} = \frac{1}{2} \left(\text{Tr}(\mathbf{Q}) + \frac{1}{\text{Tr}(\mathbf{Q})} \right) \mathbf{Q} \quad (28)$$

where $\tilde{\mathbf{K}} = \mathbf{K}/\lambda_{\max}(\mathbf{K})$. We give the following theorem.

Theorem 4: *Any $\mathbf{Q} \succeq 0$ that satisfies (27) also satisfies the conditions of Theorem 2.*

Proof: It is easy to verify that $\frac{1}{2}(\text{Tr}(\mathbf{Q}) + 1/\text{Tr}(\mathbf{Q}))\lambda_{\max}(\mathbf{K})$ is an eigenvalue of \mathbf{K} . But $\lambda_{\max}(\mathbf{K})$ is the largest eigenvalue of \mathbf{K} . On the other hand, $\frac{1}{2}(\text{Tr}(\mathbf{Q}) + 1/\text{Tr}(\mathbf{Q})) \geq 1$ holds with equality if and only if $\text{Tr}(\mathbf{Q}) = 1$. Thus, we know $\text{Tr}(\mathbf{Q}) = 1$. With this, taking trace operation over both sides of (27) leads to $\lambda_{\max}(\mathbf{K}) = \text{Tr}(\mathbf{Q}\mathbf{K})$. ■

Summarily, we can alternatively do one of the following two things to find \mathbf{Q} that satisfies the conditions of Theorem 2:

- (i) Find $\mathbf{Q} \succeq 0$ satisfies (25) and (26);
- (ii) Find $\mathbf{Q} \succeq 0$ satisfies (27);

We will discuss the algorithm to search for such \mathbf{Q} in §VII.

In the following sections, we will analyze some special cases. In particular, for $n_T = 2$ we obtain the optimal \mathbf{Q} in closed form. If $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E$ has all negative eigenvalues except a positive eigenvalue, we show that the optimal \mathbf{Q} has rank one and can also be expressed in a closed form. For general cases, we prove that the optimal \mathbf{Q} is a fixed point of a mapping from a convex set to itself, and propose an algorithm to search for it.

V. THE CASE $n_T = 2$

In this section, we analyze the case $n_T = 2$, i.e., the transmitter has two antennas. It includes the four cases $(n_T, n_R, n_E) = (2, 2, 2), (2, 2, 1), (2, 1, 2), (2, 1, 1)$. In §III, the MISO wiretap channel with $n_T = 2$ belongs to $(n_T, n_R, n_E) = (2, 1, 1)$. In [13], the case $(n_T, n_R, n_E) = (2, 2, 1)$ is analyzed. We derive the optimal \mathbf{Q} in two subsections in which $\mathbf{S}_R - \mathbf{S}_E \not\succeq 0$ or $\mathbf{S}_R - \mathbf{S}_E \succeq 0$. We also analyze the rank of optimal \mathbf{Q} with respect to SNR.

A. $\mathbf{S}_R - \mathbf{S}_E \not\succeq 0$

According to Property 3, the optimal \mathbf{Q} has low rank (rank one) and hence it has the form $\mathbf{Q} = \mathbf{u}\mathbf{u}^\dagger$ where \mathbf{u} is a unit-norm vector to be determined. We can rewrite

$$C_s(\mathbf{Q}) = \log \frac{1 + \mathbf{u}^\dagger \mathbf{S}_R \mathbf{u}}{1 + \mathbf{u}^\dagger \mathbf{S}_E \mathbf{u}} = \log \frac{\mathbf{u}^\dagger (\mathbf{I}_2 + \mathbf{S}_R) \mathbf{u}}{\mathbf{u}^\dagger (\mathbf{I}_2 + \mathbf{S}_E) \mathbf{u}}. \quad (29)$$

The optimal \mathbf{Q} is easily obtained to be $\mathbf{Q}^\circ = \mathbf{u}^\circ \mathbf{u}^{\circ \dagger}$ where \mathbf{u}° is the eigenvector associated with the largest eigenvalue of $(\mathbf{I}_2 + \mathbf{S}_E)^{-1}(\mathbf{I}_2 + \mathbf{S}_R)$. The secrecy capacity is given by

$$C_s = \log \left(\lambda_{\max} \{ (\mathbf{I}_2 + \mathbf{S}_E)^{-1} (\mathbf{I}_2 + \mathbf{S}_R) \} \right). \quad (30)$$

We can express C_s in closed form. Denote

$$\mathbf{H}_E^\dagger \mathbf{H}_E = \begin{pmatrix} a_1 & b_1 \\ b_1^* & c_1 \end{pmatrix}, \mathbf{H}_R^\dagger \mathbf{H}_R = \begin{pmatrix} a_2 & b_2 \\ b_2^* & c_2 \end{pmatrix}. \quad (31)$$

By using the fact: for any 2×2 matrix \mathbf{A} with two real eigenvalues, the largest eigenvalue is given by $\lambda_{\max}(\mathbf{A}) = [\text{Tr}(\mathbf{A}) + \sqrt{(\text{Tr}(\mathbf{A}))^2 - 4\det(\mathbf{A})}]/2$, and the matrix inverse formula

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{12}^* & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11}a_{22} - |a_{12}|^2} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{12}^* & a_{11} \end{pmatrix}, \quad (32)$$

we can obtain

$$C_s = \log \frac{A + \sqrt{A^2 - 4(B_1 - B_2)}}{2[(1/\rho + a_1)(1/\rho + c_1) - |b_1|^2]} \quad (33)$$

where $A = a_1c_2 + a_2c_1 - b_1b_2^* - b_1^*b_2 + (a_1 + a_2 + c_1 + c_2)/\rho + 2/\rho^2$, $B_1 = (a_2c_1 - b_1b_2^* + (a_2 + c_1)/\rho + 1/\rho^2)(a_1c_2 - b_1^*b_2 + (a_1 + c_2)/\rho + 1/\rho^2)$ and $B_2 = (b_2c_1 - b_1c_2 + (b_2 - b_1)/\rho)(a_1b_2^* - a_2b_1^* + (b_2^* - b_1^*)/\rho)$.

Now we analyze the secrecy degree of freedom (s.d.o.f.) defined in (14) which is different whether \mathbf{S}_E has full rank or not.

- Case 1) \mathbf{S}_E has rank two (full rank)

In this case, noting that $(\mathbf{I}_2 + \mathbf{S}_E)^{-1}(\mathbf{I}_2 + \mathbf{S}_R) = (\mathbf{I}_2/\rho + \mathbf{H}_E^\dagger \mathbf{H}_E)^{-1}(\mathbf{I}_2/\rho + \mathbf{H}_R^\dagger \mathbf{H}_R)$, we have $C_s \rightarrow \log(\lambda_{\max}\{(\mathbf{H}_E^\dagger \mathbf{H}_E)^{-1} \mathbf{H}_R^\dagger \mathbf{H}_R\})$ as $\rho \rightarrow \infty$. Thus, we get

$$s.d.o.f. = \lim_{\rho \rightarrow \infty} \frac{C_s(\rho)}{\log \rho} = 0. \quad (34)$$

- Case 2) \mathbf{S}_E has rank one (low rank)

In this case, $\mathbf{H}_E^\dagger \mathbf{H}_E$ is singular, hence can be expressed as $\mathbf{H}_E^\dagger \mathbf{H}_E = \mathbf{v}_2 \mathbf{v}_2^\dagger$. By using (60), we can write $(\mathbf{I}_2 + \rho \mathbf{S}_2)^{-1}(\mathbf{I}_2 + \rho \mathbf{S}_1) = \rho(\mathbf{I}_2 - \rho \mathbf{v}_2 \mathbf{v}_2^\dagger / (1 + \rho \|\mathbf{v}_2\|^2))(\mathbf{I}_2/\rho + \mathbf{S}_1) \rightarrow \rho(\mathbf{I}_2 - \mathbf{S}_2/\text{Tr}(\mathbf{S}_2))\mathbf{S}_1$ as $\rho \rightarrow \infty$. Thus, as $\rho \rightarrow \infty$, $\lambda_{\max}((\mathbf{I}_2 + \rho \mathbf{S}_2)^{-1}(\mathbf{I}_2 + \rho \mathbf{S}_1)) \rightarrow \rho \lambda_{\max}((\mathbf{I}_2 - \mathbf{S}_2/\text{Tr}(\mathbf{S}_2))\mathbf{S}_1)$. Thus, we get

$$s.d.o.f. = \lim_{\rho \rightarrow \infty} \frac{C_s(\rho)}{\log \rho} = 1. \quad (35)$$

We can also use (33) to obtain the same result.

B. $\mathbf{S}_R - \mathbf{S}_E \succ 0$

In this case, the optimal \mathbf{Q} may have full rank or low rank. If the optimal \mathbf{Q} has low rank, it is given in (29). Therefore, in the following we focus on the case in which the optimal \mathbf{Q} has full rank. The optimal \mathbf{Q} can be determined from the above two cases.

Since $\mathbf{Q} \succ 0$, it follows from (21) that Θ must be a positive scalar multiplication of \mathbf{I}_2 . Recall from (20) that $\Theta = \mathbf{S}_R(\mathbf{I}_2 + \mathbf{Q}\mathbf{S}_R)^{-1} - \mathbf{S}_E(\mathbf{I}_2 + \mathbf{Q}\mathbf{S}_E)^{-1}$. We know $\mathbf{S}_R \succ 0$, but \mathbf{S}_E is not necessarily positive definite. Thus, in the following, we discuss two cases respectively: a) \mathbf{S}_E has rank two (full rank); b) \mathbf{S}_E has rank one (low rank).

- Case a) \mathbf{S}_E has rank two (full rank) In this case, $\mathbf{S}_R \succ 0$, $\mathbf{S}_E \succ 0$. We can rewrite

$$\boldsymbol{\Theta} = (\mathbf{S}_R^{-1} + \mathbf{Q})^{-1} - (\mathbf{S}_E^{-1} + \mathbf{Q})^{-1} = \theta \mathbf{I}_2, \theta > 0. \quad (36)$$

Based on the eigen-decomposition $(\mathbf{S}_E^{-1} - \mathbf{S}_R^{-1})/2 = \mathbf{U}_1 \mathbf{D}_1 \mathbf{U}_1^\dagger$ where $\mathbf{D}_1 = \text{diag}(d_1, d_2)$, $d_1 \geq d_2 \geq 0$ but $d_1^2 + d_2^2 \neq 0$ (otherwise, $\mathbf{S}_R = \mathbf{S}_E$ violates the assumption $\mathbf{S}_R \neq \mathbf{S}_E$), and letting $\mathbf{C} = \mathbf{U}_1^\dagger (\mathbf{S}_E^{-1} + \mathbf{S}_R^{-1}) \mathbf{U}_1 / 2$, we get $\mathbf{S}_E^{-1} = \mathbf{U}_1 (\mathbf{C} + \mathbf{D}_1) \mathbf{U}_1^\dagger$ and $\mathbf{S}_R^{-1} = \mathbf{U}_1 (\mathbf{C} - \mathbf{D}_1) \mathbf{U}_1^\dagger$. On inserting the latter expressions in (36) we get

$$(\hat{\mathbf{Q}} + \mathbf{C} - \mathbf{D}_1)^{-1} - (\hat{\mathbf{Q}} + \mathbf{C} + \mathbf{D}_1)^{-1} = \theta \mathbf{I}_2, \theta > 0 \quad (37)$$

where $\hat{\mathbf{Q}} = \mathbf{U}_1^\dagger \mathbf{Q} \mathbf{U}_1$. Note that \mathbf{D}_1 is diagonal. We can actually show that $\hat{\mathbf{Q}} + \mathbf{C}$ must be diagonal. To prove this, let us denote the $(1, 2)$ th entry of $\hat{\mathbf{Q}} + \mathbf{C}$ by \bar{q}_{12} . We know that the $(1, 2)$ th entry of $(\hat{\mathbf{Q}} + \mathbf{C} - \mathbf{D}_1)^{-1} - (\hat{\mathbf{Q}} + \mathbf{C} + \mathbf{D}_1)^{-1}$ equals

$$-\bar{q}_{12} [\det(\hat{\mathbf{Q}} + \mathbf{C} - \mathbf{D}_1)^{-1} - \det(\hat{\mathbf{Q}} + \mathbf{C} + \mathbf{D}_1)^{-1}] = 0 \quad (38)$$

which leads to $\bar{q}_{12} = 0$. Here we used (32) and the fact $\det(\mathbf{A}) > \det(\mathbf{B})$ for $\mathbf{A} \succ \mathbf{B}$, $\mathbf{B} \succ 0$. Since $\hat{\mathbf{Q}} + \mathbf{C}$ is diagonal, we denote $\hat{\mathbf{Q}} + \mathbf{C} = \text{diag}(\bar{q}_1, \bar{q}_2)$ and $\hat{\mathbf{Q}} = \text{diag}(\bar{q}_1, \bar{q}_2) - \mathbf{C}$. It follows from $\hat{\mathbf{Q}} \succ 0$, $\text{Tr}(\hat{\mathbf{Q}}) = 1$ that $\bar{q}_1 + \bar{q}_2 = 1 + \text{Tr}(\mathbf{C})$, $\text{diag}(\bar{q}_1, \bar{q}_2) \succ \mathbf{C}$. Combining these with (37) results in

$$(d_1/d_2 - 1)\bar{q}_2^2 + 2(1 + \text{Tr}(\mathbf{C}))\bar{q}_2 + d_1(d_1 - d_2) - (1 + \text{Tr}(\mathbf{C}))^2 = 0. \quad (39)$$

We can solve \bar{q}_2 from the quadratic equation (39) and $\bar{q}_1 = 1 + \text{Tr}(\mathbf{C}) - \bar{q}_2$. If $\text{diag}(\bar{q}_1, \bar{q}_2) \succ \mathbf{C}$ holds, then $\mathbf{Q} = \mathbf{U}_1 \hat{\mathbf{Q}} \mathbf{U}_1^\dagger$ is a possible solution. If the equation (39) has no positive roots or $\text{diag}(\bar{q}_1, \bar{q}_2) \not\succ \mathbf{C}$, it means the optimal \mathbf{Q} has low rank.

- Case b) \mathbf{S}_E has rank one (low rank) In this case \mathbf{S}_E can be expressed as $\mathbf{S}_E = \mathbf{v}_e \mathbf{v}_e^\dagger$. We eigen-decompose $\mathbf{S}_E = \mathbf{U}_e \text{diag}(\lambda_e, 0) \mathbf{U}_e^\dagger$. Similarly, we get

$$\boldsymbol{\Theta} = (\mathbf{S}_R^{-1} + \mathbf{Q})^{-1} - \mathbf{S}_E (\mathbf{I}_2 + \mathbf{Q} \mathbf{S}_E)^{-1} = \theta \mathbf{I}_2, \theta > 0. \quad (40)$$

Let us define $\check{\mathbf{Q}} = \mathbf{U}_e^\dagger \mathbf{Q} \mathbf{U}_e$, $\check{\mathbf{S}}_R = \mathbf{U}_e^\dagger \mathbf{S}_R^{-1} \mathbf{U}_e$. Inserting $\check{\mathbf{Q}}$ and $\check{\mathbf{S}}_R$ into (40) results in

$$(\check{\mathbf{S}}_R + \check{\mathbf{Q}})^{-1} - \frac{\lambda_e}{1 + \check{q}_{11} \lambda_e} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \theta \mathbf{I}_2, \theta > 0. \quad (41)$$

where \check{q}_{11} is the $(1, 1)$ th entry of $\check{\mathbf{Q}}$. It follows from (41) that $\check{\mathbf{S}}_R + \check{\mathbf{Q}}$ is diagonal. Thus, we denote $\check{\mathbf{S}}_R + \check{\mathbf{Q}} = \text{diag}(\check{q}_1, \check{q}_2)$ and $\check{\mathbf{Q}} = \text{diag}(\check{q}_1, \check{q}_2) - \check{\mathbf{S}}_R$. It follows from $\check{\mathbf{Q}} \succ 0$, $\text{Tr}(\check{\mathbf{Q}}) = 1$ that $\check{q}_1 + \check{q}_2 = 1 + \text{Tr}(\check{\mathbf{S}}_R)$ and $\text{diag}(\check{q}_1, \check{q}_2) \succ \check{\mathbf{S}}_R$. Combining the above and (41) results in

$$\lambda_e \check{q}_1^2 + 2(1 - (\check{\mathbf{S}}_R)_{11} \lambda_e) \check{q}_1 - (1 - (\check{\mathbf{S}}_R)_{11} \lambda_e)(1 + \text{Tr}(\check{\mathbf{S}}_R)) = 0 \quad (42)$$

where $(\check{\mathbf{S}}_R)_{11}$ is the $(1,1)$ th entry of $\check{\mathbf{S}}_R$. We can solve \tilde{q}_1 from the quadratic equation (42), and then get $\tilde{q}_2 = 1 + \text{Tr}(\check{\mathbf{S}}_R) - \tilde{q}_1$. If $\text{diag}(\tilde{q}_1, \tilde{q}_2) \succ \check{\mathbf{S}}_R$ holds, then $\mathbf{Q} = \mathbf{U}_e \check{\mathbf{Q}} \mathbf{U}_e^\dagger$ is a possible solution. If the equation (42) has no positive roots or $\text{diag}(\tilde{q}_1, \tilde{q}_2) \not\succ \check{\mathbf{S}}_R$, it means the optimal \mathbf{Q} has low rank.

C. Rank of Optimal \mathbf{Q}

For the non-wiretap MIMO channel the rank of optimal input covariance has a non-decreasing property with respect to SNR [19]. In this section we consider the behavior of the rank of optimal input covariance of the MIMO wiretap channel with respect to SNR.

When $\mathbf{S}_R - \mathbf{S}_E \not\succ 0$, according to the result in §V-A, the optimal \mathbf{Q} has rank one, independent of SNR, and hence follows the non-decreasing property of rank. Next we focus on $\mathbf{S}_R - \mathbf{S}_E \succ 0$. According to §V-A, if the optimal \mathbf{Q} has rank one, it can be expressed as $\mathbf{Q} = \mathbf{u}_0 \mathbf{u}_0^\dagger$ where \mathbf{u}_0 is the eigenvector associated with the largest eigenvalue λ_0 of $(\mathbf{I}_2 + \rho \mathbf{S}_2)^{-1}(\mathbf{I}_2 + \rho \mathbf{S}_1)$ where $\mathbf{S}_1 = \mathbf{H}_R^\dagger \mathbf{H}_R$, $\mathbf{S}_2 = \mathbf{H}_E^\dagger \mathbf{H}_E$. Denote $\mathbf{S}_0 = \mathbf{S}_1(\mathbf{I}_2 + \rho \mathbf{u}_0 \mathbf{u}_0^\dagger \mathbf{S}_1)^{-1} - \mathbf{S}_2(\mathbf{I}_2 + \rho \mathbf{u}_0 \mathbf{u}_0^\dagger \mathbf{S}_2)^{-1}$. Since $\mathbf{S}_1 - \mathbf{S}_2 \succ 0$, we can rewrite $\mathbf{S}_0 = (\mathbf{I}_2 + \rho \mathbf{S}_1 \mathbf{u}_0 \mathbf{u}_0^\dagger)^{-1}(\mathbf{S}_1 - \mathbf{S}_2)(\mathbf{I}_2 + \rho \mathbf{u}_0 \mathbf{u}_0^\dagger \mathbf{S}_2)^{-1}$ and state that \mathbf{S}_0 has full rank (rank two). Denote

$$g(\rho) \triangleq 2\mathbf{u}_0^\dagger \mathbf{S}_0 \mathbf{u}_0 - \text{Tr}(\mathbf{S}_0) = \frac{\mathbf{u}_0^\dagger (2\mathbf{S}_1 + \rho \mathbf{S}_1^2) \mathbf{u}_0}{\mathbf{u}_0^\dagger (\mathbf{I}_{n_T} + \rho \mathbf{S}_1) \mathbf{u}_0} - \frac{\mathbf{u}_0^\dagger (2\mathbf{S}_2 + \rho \mathbf{S}_2^2) \mathbf{u}_0}{\mathbf{u}_0^\dagger (\mathbf{I}_{n_T} + \rho \mathbf{S}_2) \mathbf{u}_0} - \text{Tr}(\mathbf{S}_1 - \mathbf{S}_2) \quad (43)$$

where the matrix inverse formula (60) is used. We give the following result.

Lemma 3: *If $g(\rho) < 0$, then the optimal \mathbf{Q} has rank two; If the optimal \mathbf{Q} has rank one, there must be $g(\rho) \geq 0$.*

Please see Appendix F for details.

Then we can prove the following result.

Theorem 5: *$\lim_{\rho \rightarrow \infty} g(\rho) < 0$, hence, according to Lemma 3, there exists a certain ρ_0 such that when $\rho > \rho_0$, the optimal \mathbf{Q} has rank two.*

The proof is given in Appendix G.

Theorem 5 reveals that when the SNR is sufficient large, the optimal \mathbf{Q} always has rank two. At this point, we do not prove the rank non-decreasing property of the optimal \mathbf{Q} for the case $\mathbf{S}_R - \mathbf{S}_E \succ 0$ completely.

VI. $\mathbf{S}_R - \mathbf{S}_E$ HAS ALL NEGATIVE EIGENVALUES EXCEPT ONE POSITIVE EIGENVALUE

We analyze the case in which $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E$ has all negative eigenvalues except one positive eigenvalue, e.g., $n_T = 3$, $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E$ has two negative eigenvalues and a positive eigenvalue. In

particular, this always occurs when \mathbf{S}_E has full rank and $n_R = 1$, i.e., the legitimate receiver has a single antenna, as the following lemma stated.

Lemma 4: *Let \mathbf{x} is a known non-zero vector, \mathbf{X} is a known positive semi-definite matrix. Assume $\mathbf{x}\mathbf{x}^\dagger - \mathbf{X}$ is indefinite or positive semi-definite. If \mathbf{X} has full rank, or \mathbf{X} has rank $n_T - 1$ and \mathbf{x} is linearly independent of the eigenvectors associated with the non-zero eigenvalues of \mathbf{X} , then $\mathbf{x}\mathbf{x}^\dagger - \mathbf{X}$ has all negative eigenvalues except one positive eigenvalue.*

Proof: First, we prove the case that \mathbf{X} has full rank. Let λ be any eigenvalue of $\mathbf{x}\mathbf{x}^\dagger - \mathbf{X}$. It holds $\det(\lambda\mathbf{I} - \mathbf{x}\mathbf{x}^\dagger + \mathbf{X}) = 0$. When $\lambda \geq 0$, noting that $\lambda\mathbf{I} + \mathbf{X} \succ 0$, we get $\det(\lambda\mathbf{I} - \mathbf{x}\mathbf{x}^\dagger + \mathbf{X}) = \det(\lambda\mathbf{I} + \mathbf{X}) \det(\mathbf{I} - (\lambda\mathbf{I} + \mathbf{X})^{-1}\mathbf{x}\mathbf{x}^\dagger) = \det(\lambda\mathbf{I} + \mathbf{X})(1 - \mathbf{x}^\dagger(\lambda\mathbf{I} + \mathbf{X})^{-1}\mathbf{x}) = 0$ which leads to $\mathbf{x}^\dagger(\lambda\mathbf{I} + \mathbf{X})^{-1}\mathbf{x} = 1$. Here we use the fact $\det(\mathbf{I} + \mathbf{AB}) = \det(\mathbf{I} + \mathbf{BA})$. It is easy to prove that $\mathbf{x}^\dagger(\lambda\mathbf{I} + \mathbf{X})^{-1}\mathbf{x}$ decreases strictly with λ . Thus, there is at most one $\lambda \geq 0$ such that $\mathbf{x}^\dagger(\lambda\mathbf{I} + \mathbf{X})^{-1}\mathbf{x} = 1$. This, when combined with the fact that $\mathbf{x}\mathbf{x}^\dagger - \mathbf{X}$ is indefinite or positive semi-definite, gives the desired result.

Second, we prove the case that \mathbf{X} has rank $n_T - 1$ and \mathbf{x} is linearly independent of the eigenvectors associated with the non-zero eigenvalues of \mathbf{X} . Denote the eigen-decomposition $\mathbf{X} = \lambda_1\mathbf{u}_1\mathbf{u}_1^\dagger + \dots + \lambda_{n_T-1}\mathbf{u}_{n_T-1}\mathbf{u}_{n_T-1}^\dagger$. We can write $\mathbf{x}\mathbf{x}^\dagger - \mathbf{X} = \mathbf{F}_x\mathbf{D}_x\mathbf{F}_x^\dagger$ where $\mathbf{F}_x = [\mathbf{x}, \mathbf{u}_1, \dots, \mathbf{u}_{n_T-1}]$, $\mathbf{D}_x = \text{diag}(1, -\lambda_1, \dots, -\lambda_{n_T-1})$. Since \mathbf{x} is linearly independent of the eigenvectors associated with the non-zero eigenvalues of \mathbf{X} , it holds that \mathbf{F}_x has full rank. According to Sylvester's law of inertia [23, p.223], we know $\mathbf{F}_x\mathbf{D}_x\mathbf{F}_x^\dagger$ and \mathbf{D}_x have the same number of positive, negative, and zero eigenvalues, thus the desired result is obtained. ■

According to the above lemma: when $n_R = 1$ (hence, \mathbf{S}_R can be expressed $\mathbf{S}_R = \mathbf{v}_r\mathbf{v}_r^\dagger$) and \mathbf{S}_E has full rank, $\mathbf{S}_R - \mathbf{S}_E$ has all negative eigenvalues except one positive eigenvalue; when $n_R = 1$, \mathbf{S}_E has rank $n_T - 1$ and \mathbf{v}_r is linearly independent of the eigenvectors associated with the non-zero eigenvalues of \mathbf{S}_E , $\mathbf{S}_R - \mathbf{S}_E$ has all negative eigenvalues except one positive eigenvalue. But we point out that it does not limit to the cases in Lemma 4 in which $\mathbf{H}_R^\dagger\mathbf{H}_R - \mathbf{H}_E^\dagger\mathbf{H}_E$ has all negative eigenvalues except one positive eigenvalue. In fact, this will even occur when \mathbf{S}_R and \mathbf{S}_E both have full rank. When $\mathbf{H}_R^\dagger\mathbf{H}_R - \mathbf{H}_E^\dagger\mathbf{H}_E$ has all negative eigenvalues except one positive eigenvalue, we give the following theorem.

Theorem 6: *If $\mathbf{H}_R^\dagger\mathbf{H}_R - \mathbf{H}_E^\dagger\mathbf{H}_E$ has all negative eigenvalues except one positive eigenvalue, the optimal \mathbf{Q} has rank one. Also, the optimal \mathbf{Q} is given by $\mathbf{Q}^\circ = \mathbf{u}^\circ\mathbf{u}^{\circ\dagger}$ where \mathbf{u}° is the eigenvector associated with the largest eigenvalue of $(\mathbf{I}_{n_T} + \mathbf{S}_E)^{-1}(\mathbf{I}_{n_T} + \mathbf{S}_R)$. The secrecy capacity is given by*

$$C_s = \log(\lambda_{\max}\{(\mathbf{I}_{n_T} + \mathbf{S}_E)^{-1}(\mathbf{I}_{n_T} + \mathbf{S}_R)\}). \quad (44)$$

Please see Appendix H for details.

Similar to §V-A, if $\mathbf{H}_E^\dagger \mathbf{H}_E$ has full rank, we obtain

$$s.d.o.f = \lim_{\rho \rightarrow \infty} \frac{C_s(\rho)}{\log \rho} = 0 \quad (45)$$

and if $\mathbf{H}_E^\dagger \mathbf{H}_E$ has low rank, we obtain

$$s.d.o.f = \lim_{\rho \rightarrow \infty} \frac{C_s(\rho)}{\log \rho} = 1. \quad (46)$$

VII. ALGORITHM FOR GENERAL MIMO WIRETAP CHANNEL

In this section, we propose an algorithm to search for the optimal \mathbf{Q} which applies for any MIMO wiretap channel. The algorithm is based on the conditions of Theorem 2 or Theorem 4.

It follows from (25) that

$$\mathbf{K}^{1/2} \mathbf{Q} \mathbf{K}^{1/2} = \text{Tr}(\mathbf{Q} \mathbf{K}) \mathbf{Q} \quad (47)$$

which enables us to get

$$\mathbf{Q} = \frac{\mathbf{K}^{1/2} \mathbf{Q} \mathbf{K}^{1/2}}{\text{Tr}(\mathbf{Q} \mathbf{K})} \triangleq f(\mathbf{Q}). \quad (48)$$

Note that $f(\mathbf{Q}) \succeq 0$ and $\text{Tr}(f(\mathbf{Q})) = 1$ for any $\mathbf{Q} \in \Omega$. The equation (48) defines a mapping from a convex set to itself: $\Omega \rightarrow \Omega$, $\mathbf{Q} \mapsto f(\mathbf{Q})$. The optimal \mathbf{Q} corresponds to a fixed point of $f(\mathbf{Q})$, i.e., $f(\mathbf{Q}^\circ) = \mathbf{Q}^\circ$. To search for the fixed point, the iterative expression is

$$\mathbf{Q}^{k+1} = f(\mathbf{Q}^k), k = 0, 1, \dots \quad (49)$$

The initial point \mathbf{Q}^0 can be set to \mathbf{I}_{n_T} or choose a good initial point. The iterations stop when $\|\mathbf{Q}^{k+1} - \mathbf{Q}^k\| < 10^{-6}$. If the convergent \mathbf{Q} satisfies (26), we obtain a solution satisfying the conditions of Theorem 2, otherwise, we choose a different initial point.

VIII. NUMERICAL SIMULATIONS

We give some examples to illustrate the proposed algorithm. For illustration purpose, we consider a MIMO wiretap channel where $n_T = 4$, $n_R = 4$, $n_E = 3$.

First, we take an example for $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E \succ 0$. The channel matrices are given by

$$\mathbf{H}_R = \begin{pmatrix} -0.1107 - 0.1225i & 0.0582 - 0.3483i & 0.3239 - 0.0071i & -0.2872 - 0.2655i \\ 0.5128 - 0.3239i & -0.8903 - 0.0318i & -0.5524 - 0.0365i & -0.2072 + 0.3047i \\ -0.0041 + 0.0265i & 0.0871 - 0.0253i & 0.0183 + 1.1679i & -0.0784 + 0.0415i \\ -0.4699 - 0.1014i & -0.0888 + 0.1127i & 0.2099 + 0.3282i & 0.1734 - 0.4146i \end{pmatrix} \quad (50)$$

and

$$\mathbf{H}_E = \begin{pmatrix} -0.0766 + 0.1370i & -0.0977 - 0.0985i & 0.0002 - 0.0695i & 0.0583 + 0.0356i \\ -0.0355 - 0.1167i & 0.1607 - 0.1091i & -0.0809 + 0.1481i & -0.0218 + 0.1109i \\ 0.1375 - 0.0381i & -0.0845 - 0.0610i & -0.0011 + 0.1129i & -0.0393 + 0.1124i \end{pmatrix}. \quad (51)$$

The eigenvalues of $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E$ are 0.0085, 0.3704, 0.8945, 2.5213. Fig. 1-3 depict respectively the eigenvalues of \mathbf{Q}^k , secrecy rate and $\|\mathbf{Q}^{k+1} - \mathbf{Q}^k\|$ in the iterations where the SNR is $\rho = 8$ dB. Fig. 4-5 depict respectively the (possible) secrecy capacity and eigenvalues of (possible) optimal \mathbf{Q} under different SNRs. It can be seen from Fig. 5 that the (possible) optimal \mathbf{Q} can have rank one to four with the increasing SNR, which shows that when $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E \succ 0$, the (possible) optimal \mathbf{Q} may have low rank or full rank.

Secondly, we take an example for $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E \not\succ 0$. The channel matrices are given by

$$\mathbf{H}_R = \begin{pmatrix} -0.1110 - 0.0667i & -0.1937 - 0.1349i & -0.0752 - 0.2707i & -0.2718 + 0.2730i \\ 0.2877 + 0.6779i & -0.7832 - 0.2249i & 0.4350 + 0.2637i & 0.4160 + 0.5109i \\ 0.3266 - 0.2779i & -0.2345 - 0.4472i & 0.2448 + 0.3488i & -0.6794 - 0.0117i \\ -0.1221 + 0.4915i & 0.0959 - 0.2557i & -0.0219 + 0.5077i & 0.1449 + 0.3294i \end{pmatrix} \quad (52)$$

and

$$\mathbf{H}_E = \begin{pmatrix} 0.1468 - 0.1185i & 0.4071 + 0.4469i & 0.2474 - 0.3291i & -0.6264 - 0.1313i \\ -0.0520 + 0.2917i & -0.4978 + 0.0545i & 0.0779 - 0.3472i & -0.0132 - 0.1327i \\ 0.5799 - 0.1767i & 0.2298 + 0.3331i & -0.1151 - 0.2000i & 0.1404 - 0.3501i \end{pmatrix}. \quad (53)$$

The eigenvalues of $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E$ are $-0.8206, -0.1565, 0.9365, 1.8506$. Figures 6-8 depict respectively the eigenvalues of \mathbf{Q}^k , secrecy rate and $\|\mathbf{Q}^{k+1} - \mathbf{Q}^k\|$ in the iterations where the SNR is $\rho = 8$ dB. Figures 9-10 depict respectively the (possible) secrecy capacity and eigenvalues of (possible) optimal \mathbf{Q} under different SNRs. It can be seen from Fig. 10 that the (possible) optimal \mathbf{Q} always has rank two, which equals the number of positive eigenvalues of $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E$.

IX. CONCLUSION

We have investigated the problem of finding the optimal input covariance matrix that achieves secrecy capacity subject to a power constraint. In particular, for the multiple-input single-output (MISO) wiretap channel, the optimal input covariance matrix is obtained in closed form. For general cases, we derive the necessary conditions for the optimal solution consisting of a set of equations. For the case in which the

transmitter has two antennas, the derived necessary conditions can result in a closed form solution. If the difference is indefinite and has all negative eigenvalues except one positive eigenvalue, we prove that the optimal input covariance matrix has rank one and can be obtained in closed form. For other cases, we prove that the solution is a fixed point of a mapping from a convex set to itself and provide an iterative procedure to search for it.

APPENDIX A

PROOF OF LEMMA 1

First we show the necessary part. When $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E$ is negative semi-definite, so is $\mathbf{R}_x^{1/2}(\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E)\mathbf{R}_x^{1/2}$ which leads to $\mathbf{I}_{n_T} + \mathbf{R}_x^{1/2} \mathbf{H}_E^\dagger \mathbf{H}_E \mathbf{R}_x^{1/2} / \sigma^2 \succeq \mathbf{I}_{n_T} + \mathbf{R}_x^{1/2} \mathbf{H}_R^\dagger \mathbf{H}_R \mathbf{R}_x^{1/2} / \sigma^2$. With this, using the fact: if $\mathbf{A} \succ 0$, $\mathbf{B} \succ 0$, $\mathbf{A} \succeq \mathbf{B}$ then $\det(\mathbf{A}) \geq \det(\mathbf{B})$, and applying the identity $\det(\mathbf{I} + \mathbf{AB}) = \det(\mathbf{I} + \mathbf{BA})$ to (3) results in $C_s(\mathbf{R}_x) \leq 0$. To show the sufficient part, we rewrite (3) as

$$C_s(\mathbf{R}_x) = \log \det \left(\mathbf{I}_{n_T} + \frac{1}{\sigma^2} \mathbf{R}_x^{1/2} (\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E) \mathbf{R}_x^{1/2} (\mathbf{I}_{n_T} + \mathbf{R}_x^{1/2} \mathbf{H}_E^\dagger \mathbf{H}_E \mathbf{R}_x^{1/2} / \sigma^2)^{-1} \right). \quad (54)$$

Note that $(\mathbf{I}_{n_T} + \mathbf{R}_x^{1/2} \mathbf{H}_E^\dagger \mathbf{H}_E \mathbf{R}_x^{1/2} / \sigma^2)^{-1} \succ 0$, and it suffices to show that there exists \mathbf{R}_x such that $\mathbf{R}_x^{1/2} (\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E) \mathbf{R}_x^{1/2} \succeq 0$. Let us define the eigen-decomposition $\mathbf{H}_R^\dagger \mathbf{H}_R - \mathbf{H}_E^\dagger \mathbf{H}_E = \mathbf{U}_r \mathbf{D}_r \mathbf{U}_r^\dagger$. It is easy to verify that $\mathbf{R}_x = \mathbf{U}_r \mathbf{D}_x \mathbf{U}_r^\dagger$ is a choice where the entries of the diagonal \mathbf{D}_x are zero corresponding to the position of negative entries in \mathbf{D}_r .

APPENDIX B

PROOF OF PROPERTY 1

First, we prove the former part. If $\mathbf{S}_R \succ 0$ and $\mathbf{S}_E \succ 0$, then we can rewrite (20) as

$$\mathbf{\Theta} = (\mathbf{S}_R^{-1} + \mathbf{Q})^{-1} - (\mathbf{S}_E^{-1} + \mathbf{Q})^{-1}. \quad (55)$$

We can state that $\mathbf{\Theta}$ is not negative semi-definite, otherwise, we get $\mathbf{S}_E \succeq \mathbf{S}_R$ which violates the assumption that $\mathbf{S}_R - \mathbf{S}_E$ is indefinite or positive semi-definite.

Next we consider the case that \mathbf{S}_R or \mathbf{S}_E are singular. Denote $\mathbf{S}_R - \mathbf{S}_E = \mathbf{\Delta}$. By using the fact: $\mathbf{S}_E(\mathbf{I} + \mathbf{QS}_E)^{-1} = (\mathbf{I} + \mathbf{S}_E\mathbf{Q})^{-1}\mathbf{S}_E$, we can rewrite (20) as

$$\begin{aligned} \mathbf{\Theta} &= \mathbf{S}_R(\mathbf{I}_{n_T} + \mathbf{QS}_R)^{-1} - (\mathbf{I}_{n_T} + \mathbf{S}_E\mathbf{Q})^{-1}\mathbf{S}_E \\ &= (\mathbf{I}_{n_T} + \mathbf{S}_E\mathbf{Q})^{-1}\mathbf{\Delta}(\mathbf{I}_{n_T} + \mathbf{QS}_R)^{-1} \\ &= (\mathbf{I}_{n_T} + \mathbf{S}_E\mathbf{Q})^{-1}\mathbf{P}(\mathbf{I}_{n_T} + \mathbf{QS}_E)^{-1} \end{aligned} \quad (56)$$

where $\mathbf{P} = \mathbf{\Delta}(\mathbf{I}_{n_T} + (\mathbf{I}_{n_T} + \mathbf{Q}\mathbf{S}_E)^{-1}\mathbf{Q}\mathbf{\Delta})^{-1}$. According to Sylvester's law of inertia [23, p.223], it suffices to prove that \mathbf{P} has positive eigenvalue. If $\mathbf{\Delta}$ is nonsingular, we can write

$$\mathbf{P} = (\mathbf{\Delta}^{-1} + (\mathbf{I}_{n_T} + \mathbf{Q}\mathbf{S}_E)^{-1}\mathbf{Q})^{-1}. \quad (57)$$

Note that $(\mathbf{I}_{n_T} + \mathbf{Q}\mathbf{S}_E)^{-1}\mathbf{Q} \succeq 0$ and using the assumption $\mathbf{\Delta}$ is indefinite or positive semi-definite, hence, we get that \mathbf{P} has positive eigenvalue. If $\mathbf{\Delta}$ is singular, there exists $\delta_0 > 0$ such that for any $0 < \delta < \delta_0$, $\mathbf{\Delta}' = \mathbf{\Delta} - \delta\mathbf{I}_{n_T}$ is nonsingular and also indefinite or positive semi-definite. Similarly, we can prove that $\mathbf{P}' = \mathbf{\Delta}'(\mathbf{I}_{n_T} + (\mathbf{I}_{n_T} + \mathbf{Q}\mathbf{S}_E)^{-1}\mathbf{Q}\mathbf{\Delta}')^{-1}$ has positive eigenvalue. Next we prove $\mathbf{P} \succ \mathbf{P}'$. Denote $\mathbf{W} = (\mathbf{I}_{n_T} + \mathbf{Q}\mathbf{S}_E)^{-1}\mathbf{Q}$. Similar to the skill in (56), we get

$$\mathbf{P} - \mathbf{P}' = (\mathbf{I}_{n_T} + \mathbf{\Delta}\mathbf{W})^{-1}\mathbf{W}_1(\mathbf{I}_{n_T} + \mathbf{W}\mathbf{\Delta})^{-1} \quad (58)$$

where $\mathbf{W}_1 = \delta(\mathbf{I}_{n_T} - \delta(\mathbf{I}_{n_T} + \mathbf{W}\mathbf{\Delta})^{-1}\mathbf{W})^{-1}$. We can see that when δ is sufficient small, \mathbf{W}_1 is positive definite. Thus, $\mathbf{P} \succ \mathbf{P}'$ and \mathbf{P} has positive eigenvalue.

Second, we prove the latter part. Denote the eigen-decomposition $\mathbf{\Theta} = \mathbf{U}_\Theta \mathbf{D}_\Theta \mathbf{U}_\Theta^\dagger$ and let $\mathbf{Q}_1 = \mathbf{U}_\Theta^\dagger \mathbf{Q} \mathbf{U}_\Theta$. We know $\mathbf{Q}_1 \succeq 0$ and $\text{Tr}(\mathbf{Q}_1) = \text{Tr}(\mathbf{Q})$. With these, we can write

$$\text{Tr}(\mathbf{Q}\mathbf{\Theta}) = \text{Tr}(\mathbf{Q}_1 \mathbf{D}_\Theta) = \text{Tr}(\text{diag}(\mathbf{Q}_1) \mathbf{D}_\Theta) \leq \text{Tr}(\mathbf{Q}_1) \lambda_{\max}(\mathbf{D}_\Theta) = \text{Tr}(\mathbf{Q}) \lambda_{\max}(\mathbf{\Theta}). \quad (59)$$

In particular, if $\text{Tr}(\mathbf{Q}) = 1$, we get $\text{Tr}(\mathbf{Q}\mathbf{\Theta}) \leq \lambda_{\max}(\mathbf{\Theta})$.

APPENDIX C

PROOF OF THEOREM 2

It follows from (17) that $\mathbf{\Psi}\mathbf{Q} = \mathbf{Q}\mathbf{\Psi} = 0$, that is, $\mathbf{\Psi}$ and \mathbf{Q} commute and have the same eigenvectors [22, p.239] and their eigenvalue patterns are complementary in the sense that if $\lambda_i(\mathbf{Q}) > 0$, then $\lambda_i(\mathbf{\Psi}) = 0$, and vice versa [17]. This result, when combined with (16), implies that $\mathbf{\Theta}$ and \mathbf{Q} commute and have the same eigenvectors, i.e., they have the eigen-decompositions $\mathbf{Q} = \mathbf{U}_q \mathbf{D}_q \mathbf{U}_q^\dagger$ and $\mathbf{\Theta} = \mathbf{U}_q \mathbf{D}_\Theta \mathbf{U}_q^\dagger$. Further, we get $\mathbf{\Theta}\mathbf{Q} = \mathbf{Q}\mathbf{\Theta} = \theta\mathbf{Q}$, which, when combined with $\text{Tr}(\mathbf{Q}) = 1$ and the fact $\text{Tr}(\mathbf{Q}\mathbf{\Theta}) = \text{Tr}(\mathbf{Q}^{1/2}\mathbf{\Theta}\mathbf{Q}^{1/2})$ is always real, leads to $\theta = \text{Tr}(\mathbf{Q}\mathbf{\Theta})$ and (21) (also see [18]).

The condition (21) reveals that for the optimal \mathbf{Q} , $\mathbf{Q}\mathbf{\Theta}$ is a scaled version of \mathbf{Q} . Further the eigenvalues of $\mathbf{\Theta}$ corresponding to the positive eigenvalues of \mathbf{Q} are all equal to $\text{Tr}(\mathbf{Q}\mathbf{\Theta})$, while the remaining eigenvalues of $\mathbf{\Theta}$ are all less than or equal to $\text{Tr}(\mathbf{Q}\mathbf{\Theta})$, which follows from (16), (21) and $\mathbf{\Psi} \succeq 0$. Based on the above, it holds the second condition (22).

APPENDIX D

PROOF OF THEOREM 3

In the MISO wiretap channel, $\mathbf{S}_R = \rho \mathbf{h}_R \mathbf{h}_R^\dagger$ and $\mathbf{S}_E = \rho \mathbf{h}_E \mathbf{h}_E^\dagger$. By using the matrix inverse formula for two vectors \mathbf{x} and \mathbf{y}

$$(\mathbf{I} + \mathbf{x} \mathbf{y}^\dagger)^{-1} = \mathbf{I} - \mathbf{x} \mathbf{y}^\dagger / (1 + \mathbf{y}^\dagger \mathbf{x}), \quad (60)$$

We can write

$$\mathbf{\Theta} = \frac{\rho \mathbf{h}_R \mathbf{h}_R^\dagger}{1 + \rho \mathbf{h}_R^\dagger \mathbf{Q} \mathbf{h}_R} - \frac{\rho \mathbf{h}_E \mathbf{h}_E^\dagger}{1 + \rho \mathbf{h}_E^\dagger \mathbf{Q} \mathbf{h}_E}. \quad (61)$$

That is to say, $\mathbf{\Theta}$ has the form of $\alpha_1 \mathbf{h}_R \mathbf{h}_R^\dagger - \alpha_2 \mathbf{h}_E \mathbf{h}_E^\dagger$, $\alpha_1 > 0, \alpha_2 > 0$. According to Lemma 2, we know: if $\mathbf{h}_R = \xi \mathbf{h}_E$, then $\mathbf{\Theta}$ has only one nonzero eigenvalue; if $\mathbf{h}_R \neq \mathbf{h}_E$, then $\mathbf{\Theta}$ has only two nonzero eigenvalues, one is positive and the other is negative. With this, since \mathbf{Q} satisfies (21), it is easy to verify \mathbf{Q} has rank one. Let $\mathbf{Q} = \mathbf{u} \mathbf{u}^\dagger$ and we have

$$\mathbf{\Theta} = \frac{\rho \mathbf{h}_R \mathbf{h}_R^\dagger}{1 + \rho \mathbf{h}_R^\dagger \mathbf{u} \mathbf{u}^\dagger \mathbf{h}_R} - \frac{\rho \mathbf{h}_E \mathbf{h}_E^\dagger}{1 + \rho \mathbf{h}_E^\dagger \mathbf{u} \mathbf{u}^\dagger \mathbf{h}_E} \quad (62)$$

$$\text{Tr}(\mathbf{Q} \mathbf{\Theta}) = \frac{1}{1 + \rho \mathbf{h}_E^\dagger \mathbf{u} \mathbf{u}^\dagger \mathbf{h}_E} - \frac{1}{1 + \rho \mathbf{h}_R^\dagger \mathbf{u} \mathbf{u}^\dagger \mathbf{h}_R}. \quad (63)$$

Let $\omega_1 = 1 + \rho \mathbf{h}_R^\dagger \mathbf{u} \mathbf{u}^\dagger \mathbf{h}_R$, $\omega_2 = 1 + \rho \mathbf{h}_E^\dagger \mathbf{u} \mathbf{u}^\dagger \mathbf{h}_E$. According to Lemma 2, the largest eigenvalue of $\mathbf{\Theta}$ is given by

$$\lambda_{\max}(\mathbf{\Theta}) = \frac{\rho \|\mathbf{h}_R\|^2}{2\omega_1} - \frac{\rho \|\mathbf{h}_E\|^2}{2\omega_2} + \frac{1}{2} \sqrt{\left(\frac{\rho \|\mathbf{h}_R\|^2}{\omega_1} + \frac{\rho \|\mathbf{h}_E\|^2}{\omega_2} \right)^2 - \frac{4\rho^2 |\mathbf{h}_R^\dagger \mathbf{h}_E|^2}{\omega_1 \omega_2}}. \quad (64)$$

Since \mathbf{Q} satisfies (22), we have

$$\frac{\rho \|\mathbf{h}_R\|^2}{2\omega_1} - \frac{\rho \|\mathbf{h}_E\|^2}{2\omega_2} + \frac{1}{2} \sqrt{\left(\frac{\rho \|\mathbf{h}_R\|^2}{\omega_1} + \frac{\rho \|\mathbf{h}_E\|^2}{\omega_2} \right)^2 - \frac{4\rho^2 |\mathbf{h}_R^\dagger \mathbf{h}_E|^2}{\omega_1 \omega_2}} = \frac{1}{\omega_2} - \frac{1}{\omega_1} \quad (65)$$

which leads to

$$1 + \frac{\rho \|\mathbf{h}_R\|^2}{2} - \left(1 + \frac{\rho \|\mathbf{h}_E\|^2}{2} \right) \frac{\omega_1}{\omega_2} + \frac{\rho}{2} \sqrt{\left(\|\mathbf{h}_R\|^2 + \|\mathbf{h}_E\|^2 \frac{\omega_1}{\omega_2} \right)^2 - 4 |\mathbf{h}_R^\dagger \mathbf{h}_E|^2 \frac{\omega_1}{\omega_2}} = 0. \quad (66)$$

This equation (66) is exactly (11), i.e., $F(\alpha) = 0$ where $\alpha = \omega_1/\omega_2$. On the other hand, we know

$$\frac{\omega_1}{\omega_2} = \frac{1 + \rho \mathbf{h}_R^\dagger \mathbf{u} \mathbf{u}^\dagger \mathbf{h}_R}{1 + \rho \mathbf{h}_E^\dagger \mathbf{u} \mathbf{u}^\dagger \mathbf{h}_E}. \quad (67)$$

According to the result in §III, the root of $F(\alpha) = 0$ corresponds to the maximization of the right hand side (RHS) of (67) (see also (7)). Thus, the conditions of Theorem 2 guarantee the optimal input covariance.

APPENDIX E
PROOF OF PROPERTY 2

From (20), we know that $\mathbf{Q}\Theta = (\mathbf{I}_{n_T} + \mathbf{Q}\mathbf{S}_E)^{-1} - (\mathbf{I}_{n_T} + \mathbf{Q}\mathbf{S}_R)^{-1}$. With this, Left-multiplication by $\mathbf{I}_{n_T} + \mathbf{Q}\mathbf{S}_E$ and right-multiplication by $(\mathbf{I}_{n_T} + \mathbf{Q}\mathbf{S}_R)\mathbf{Q}$ of both sides of (21) results in

$$\mathbf{Q}(\mathbf{S}_R - \mathbf{S}_E)\mathbf{Q} = \text{Tr}(\mathbf{Q}\Theta)(\mathbf{Q} + \mathbf{Q}\mathbf{S}_E\mathbf{Q})(\mathbf{Q} + \mathbf{Q}\mathbf{S}_R\mathbf{Q}). \quad (68)$$

Note that the left hand side (LHS) of (68) is Hermitian, hence, $(\mathbf{Q} + \mathbf{Q}\mathbf{S}_E\mathbf{Q})(\mathbf{Q} + \mathbf{Q}\mathbf{S}_R\mathbf{Q})$ is Hermitian as well, which implies that the matrices $\mathbf{Q} + \mathbf{Q}\mathbf{S}_E\mathbf{Q}$ and $\mathbf{Q} + \mathbf{Q}\mathbf{S}_R\mathbf{Q}$ commute and have the same eigenvectors. On the other hand, the RHS of (68) has all non-negative eigenvalues, thus $\mathbf{Q}(\mathbf{S}_R - \mathbf{S}_E)\mathbf{Q} \succeq 0$.

Finally, recalling (56), $\Theta = (\mathbf{I}_{n_T} + \mathbf{S}_E\mathbf{Q})^{-1}(\mathbf{S}_R - \mathbf{S}_E)(\mathbf{I}_{n_T} + \mathbf{Q}\mathbf{S}_R)^{-1}$, we know $\text{rank}(\Theta) = \text{rank}(\mathbf{S}_R - \mathbf{S}_E)$ which follows from the fact $\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}) - n \leq \text{rank}(\mathbf{AB}) \leq \min\{\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\}$ for two $n \times n$ matrices \mathbf{A} and \mathbf{B} [23]. Further, since $\text{rank}(\mathbf{Q}\Theta) \leq \min\{\text{rank}(\mathbf{Q}), \text{rank}(\Theta)\}$, if $\text{rank}(\Theta) < \text{rank}(\mathbf{Q})$, then $\text{rank}(\mathbf{Q}\Theta) < \text{rank}(\mathbf{Q})$, but this violates (21). Thus, $\text{rank}(\Theta) \geq \text{rank}(\mathbf{Q})$.

APPENDIX F
PROOF OF LEMMA 3

First, we can prove that \mathbf{u}_0 is the eigenvector of \mathbf{S}_0 . To prove this, it suffices to show that $\mathbf{S}_0\mathbf{u}_0 = \eta_0\mathbf{u}_0$ for a certain scalar η_0 . By using the formula (60), we can write

$$\mathbf{S}_0\mathbf{u}_0 = \frac{\mathbf{S}_1\mathbf{u}_0}{1 + \rho\mathbf{u}_0^\dagger\mathbf{S}_1\mathbf{u}_0} - \frac{\mathbf{S}_2\mathbf{u}_0}{1 + \rho\mathbf{u}_0^\dagger\mathbf{S}_2\mathbf{u}_0}. \quad (69)$$

Since \mathbf{u}_0 is the eigenvector of $(\mathbf{I}_2 + \rho\mathbf{S}_2)^{-1}(\mathbf{I}_2 + \rho\mathbf{S}_1)$ associated with the eigenvalue λ_0 , it holds $(\mathbf{I}_2 + \rho\mathbf{S}_1)\mathbf{u}_0 = \lambda_0(\mathbf{I}_2 + \rho\mathbf{S}_2)\mathbf{u}_0$ which leads to two facts: $1 + \rho\mathbf{u}_0^\dagger\mathbf{S}_1\mathbf{u}_0 = \lambda_0(1 + \rho\mathbf{u}_0^\dagger\mathbf{S}_2\mathbf{u}_0)$ and $(\mathbf{S}_1 - \lambda_0\mathbf{S}_2)\mathbf{u}_0 = \frac{\lambda_0 - 1}{\rho}\mathbf{u}_0$. Inserting them into (69) gives $\mathbf{S}_0\mathbf{u}_0 = \eta_0\mathbf{u}_0$ where $\eta_0 = (\lambda_0 - 1)/(\rho(1 + \rho\mathbf{u}_0^\dagger\mathbf{S}_1\mathbf{u}_0))$. We also know $\eta_0 = \mathbf{u}_0^\dagger\mathbf{S}_0\mathbf{u}_0$.

Second, let \mathbf{v} is any unit-norm vector and $\mathbf{v} \neq \xi\mathbf{u}_0$. Define the function

$$\begin{aligned} g(t, \mathbf{v}) &= \log \det(\mathbf{I}_2 + \rho((1-t)\mathbf{u}_0\mathbf{u}_0^\dagger + t\mathbf{v}\mathbf{v}^\dagger)\mathbf{S}_1) \\ &\quad - \log \det(\mathbf{I}_2 + \rho((1-t)\mathbf{u}_0\mathbf{u}_0^\dagger + t\mathbf{v}\mathbf{v}^\dagger)\mathbf{S}_2), \quad t \in [0, 1]. \end{aligned} \quad (70)$$

It is easy to know: $(1-t)\mathbf{u}_0\mathbf{u}_0^\dagger + t\mathbf{v}\mathbf{v}^\dagger \in \Omega$; when $t = 0$, $(1-t)\mathbf{u}_0\mathbf{u}_0^\dagger + t\mathbf{v}\mathbf{v}^\dagger = \mathbf{u}_0\mathbf{u}_0^\dagger$; when $t > 0$, $(1-t)\mathbf{u}_0\mathbf{u}_0^\dagger + t\mathbf{v}\mathbf{v}^\dagger$ has rank two. We can state that if $\frac{\partial g}{\partial t}|_{t=0} > 0$ for a certain \mathbf{v}_0 , the optimal \mathbf{Q} has rank two. The reason is simple: assume that the optimal \mathbf{Q} has rank one, then it holds $g(0, \mathbf{v}) \geq g(t, \mathbf{v})$, $\forall t, \mathbf{v}$.

But on the other hand, it follows from $\frac{\partial g}{\partial t}|_{t=0} > 0$ that there exists $t_0 > 0$ such that $g(t_0, \mathbf{v}_0) > g(0, \mathbf{v}_0)$. This produces a contradiction. By using the derivative formula [16]

$$\frac{\partial \log \det(\mathbf{A} + t\mathbf{B})}{\partial t} = \text{Tr}\{\mathbf{B}(\mathbf{A} + t\mathbf{B})^{-1}\}, \quad (71)$$

we can obtain

$$\frac{\partial g}{\partial t}|_{t=0} = \text{Tr}\{\rho(\mathbf{v}\mathbf{v}^\dagger - \mathbf{u}_0\mathbf{u}_0^\dagger)\mathbf{S}_0\} = \rho(\mathbf{v}^\dagger\mathbf{S}_0\mathbf{v} - \mathbf{u}_0^\dagger\mathbf{S}_0\mathbf{u}_0). \quad (72)$$

It follows from (72) that if $\mathbf{u}_0^\dagger\mathbf{S}_0\mathbf{u}_0 < \lambda_{\max}(\mathbf{S}_0)$, there always exists \mathbf{v} such that $\frac{\partial g}{\partial t}|_{t=0} > 0$, and hence the optimal \mathbf{Q} has rank two; if the optimal \mathbf{Q} has rank one, there must be $\frac{\partial g}{\partial t}|_{t=0} \leq 0$ which leads to $\mathbf{u}_0^\dagger\mathbf{S}_0\mathbf{u}_0 = \lambda_{\max}(\mathbf{S}_0)$. Further, since \mathbf{u}_0 is the eigenvector of \mathbf{S}_0 , thus $\mathbf{u}_0^\dagger\mathbf{S}_0\mathbf{u}_0 < \lambda_{\max}(\mathbf{S}_0)$ and $\mathbf{u}_0^\dagger\mathbf{S}_0\mathbf{u}_0 = \lambda_{\max}(\mathbf{S}_0)$ are equivalent to $2\mathbf{u}_0^\dagger\mathbf{S}_0\mathbf{u}_0 < \text{Tr}(\mathbf{S}_0)$ and $2\mathbf{u}_0^\dagger\mathbf{S}_0\mathbf{u}_0 \geq \text{Tr}(\mathbf{S}_0)$, respectively. This completes the proof.

APPENDIX G

PROOF OF THEOREM 5

We discuss two cases respectively.

- (i) \mathbf{S}_2 has full rank. Since $(\mathbf{I}_2 + \rho\mathbf{S}_2)^{-1}(\mathbf{I}_2 + \rho\mathbf{S}_1) = (\mathbf{I}_2/\rho + \mathbf{S}_2)^{-1}(\mathbf{I}_2/\rho + \mathbf{S}_1)$, we know that as $\rho \rightarrow \infty$, \mathbf{u}_0 goes to the eigenvector $\tilde{\mathbf{u}}_0$ associated with the largest eigenvalue λ of $\mathbf{S}_2^{-1}\mathbf{S}_1$. We can write $\mathbf{S}_1\tilde{\mathbf{u}}_0 = \lambda\mathbf{S}_2\tilde{\mathbf{u}}_0$, hence $\tilde{\mathbf{u}}_0^\dagger\mathbf{S}_1^2\tilde{\mathbf{u}}_0 = \lambda\tilde{\mathbf{u}}_0^\dagger\mathbf{S}_1\mathbf{S}_2\tilde{\mathbf{u}}_0$, $\tilde{\mathbf{u}}_0^\dagger\mathbf{S}_1\tilde{\mathbf{u}}_0 = \lambda\tilde{\mathbf{u}}_0^\dagger\mathbf{S}_2\tilde{\mathbf{u}}_0$. Then

$$\begin{aligned} g(\rho) &\rightarrow \frac{\tilde{\mathbf{u}}_0^\dagger\mathbf{S}_1^2\tilde{\mathbf{u}}_0}{\tilde{\mathbf{u}}_0^\dagger\mathbf{S}_1\tilde{\mathbf{u}}_0} - \frac{\tilde{\mathbf{u}}_0^\dagger\mathbf{S}_2^2\tilde{\mathbf{u}}_0}{\tilde{\mathbf{u}}_0^\dagger\mathbf{S}_2\tilde{\mathbf{u}}_0} - \text{Tr}(\mathbf{S}_1 - \mathbf{S}_2) = \frac{\tilde{\mathbf{u}}_0^\dagger(\mathbf{S}_1 - \mathbf{S}_2)\mathbf{S}_2\tilde{\mathbf{u}}_0}{\tilde{\mathbf{u}}_0^\dagger\mathbf{S}_2\tilde{\mathbf{u}}_0} - \text{Tr}(\mathbf{S}_1 - \mathbf{S}_2) \\ &\leq \lambda_{\max}(\mathbf{S}_1 - \mathbf{S}_2) - \text{Tr}(\mathbf{S}_1 - \mathbf{S}_2) < 0 \end{aligned} \quad (73)$$

where we use the fact $\mathbf{x}^\dagger\mathbf{A}\mathbf{x}/(\mathbf{x}^\dagger\mathbf{B}\mathbf{x}) \leq \lambda_{\max}(\mathbf{A}\mathbf{B}^{-1})$.

- (ii) \mathbf{S}_2 has rank one and hence can be expressed as $\mathbf{S}_2 = \mathbf{v}_2\mathbf{v}_2^\dagger$. As $\rho \rightarrow \infty$,

$$\begin{aligned} g(\rho) &\rightarrow \frac{\mathbf{u}_0^\dagger\mathbf{S}_1^2\mathbf{u}_0}{\mathbf{u}_0^\dagger\mathbf{S}_1\mathbf{u}_0} - \frac{\mathbf{u}_0^\dagger\mathbf{S}_2^2\mathbf{u}_0}{\mathbf{u}_0^\dagger\mathbf{S}_2\mathbf{u}_0} - \text{Tr}(\mathbf{S}_1 - \mathbf{S}_2) = \frac{\mathbf{u}_0^\dagger\mathbf{S}_1^2\mathbf{u}_0}{\mathbf{u}_0^\dagger\mathbf{S}_1\mathbf{u}_0} - \text{Tr}(\mathbf{S}_2) - \text{Tr}(\mathbf{S}_1 - \mathbf{S}_2) \\ &\leq \lambda_{\max}(\mathbf{S}_1) - \text{Tr}(\mathbf{S}_1) < 0. \end{aligned} \quad (74)$$

APPENDIX H

PROOF OF THEOREM 6

According to Property 2, we know $\mathbf{Q}(\mathbf{S}_R - \mathbf{S}_E)\mathbf{Q} \succeq 0$. We will show that under the condition of Theorem 6, any $\mathbf{Q} \succeq 0$ that satisfies $\mathbf{Q}(\mathbf{S}_R - \mathbf{S}_E)\mathbf{Q} \succeq 0$ has rank one. To prove this, we use the

mathematical induction which consists of two steps: (1) showing that the statement holds when $n_T = 2$; (2) showing that if the statement holds for some $n_T = k \geq 2$, then the statement also holds when $k + 1$ is substituted for k .

First, we address (1). It follows from Property 3 that the optimal \mathbf{Q} has low rank, i.e., rank one. Next we deal with (2). We denote the eigen-decomposition $\mathbf{S}_R - \mathbf{S}_E = \mathbf{U}\mathbf{D}\mathbf{U}^\dagger$ where \mathbf{D} is a diagonal matrix with all negative diagonal entries except a positive one. Let $\mathbf{X} = \mathbf{U}^\dagger \mathbf{Q} \mathbf{U}$. Then $\mathbf{Q}(\mathbf{S}_R - \mathbf{S}_E)\mathbf{Q} \succeq 0$ is equivalent to $\mathbf{X}\mathbf{D}\mathbf{X} \succeq 0$. It suffices to prove the following:

Problem: Assume the following is true: any $k \times k$ diagonal matrix \mathbf{D} which has $k - 1$ negative entries and one positive one in its diagonal, then any $k \times k$ matrix $\mathbf{X} \succeq 0$ that satisfies $\mathbf{X}\mathbf{D}\mathbf{X} \succeq 0$ always has rank one. Is it true for $k + 1$?

Denote the $(k + 1) \times (k + 1)$ diagonal matrix

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0}^\dagger & -\xi \end{pmatrix} \quad (75)$$

where \mathbf{D}_1 is a $k \times k$ diagonal matrix with all negative diagonal entries except a positive one, $\xi > 0$. We also denote the $(k + 1) \times (k + 1)$ positive semi-definite matrix

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{b} \\ \mathbf{b}^\dagger & x \end{pmatrix} \quad (76)$$

where \mathbf{X}_1 is a $k \times k$ matrix, \mathbf{b} is a $k \times 1$ vector, x is a scalar. It follows from $\mathbf{X} \succeq 0$ that: $\mathbf{X}_1 \succeq 0$, $x \geq 0$; if $x = 0$, then $\mathbf{b} = \mathbf{0}$ [23]. Now we can write

$$\mathbf{X}\mathbf{D}\mathbf{X} = \begin{pmatrix} \mathbf{X}_1\mathbf{D}_1\mathbf{X}_1 - \xi\mathbf{b}\mathbf{b}^\dagger & \mathbf{X}_1\mathbf{D}_1\mathbf{b} - \xi x\mathbf{b} \\ \mathbf{b}^\dagger\mathbf{D}_1\mathbf{X}_1 - \xi x\mathbf{b}^\dagger & \mathbf{b}^\dagger\mathbf{D}_1\mathbf{b} - \xi x^2 \end{pmatrix}. \quad (77)$$

When $x = 0$, there will be $\mathbf{b} = \mathbf{0}$ and hence $\mathbf{X}\mathbf{D}\mathbf{X} = \text{diag}(\mathbf{X}_1\mathbf{D}_1\mathbf{X}_1, 0) \succeq 0$ which is equivalent to $\mathbf{X}_1\mathbf{D}_1\mathbf{X}_1 \succeq 0$. With this, based on the assumption for k , \mathbf{X}_1 has rank one, thus, \mathbf{X} has rank one. In the following, we consider $x > 0$. Combining (77) with $\mathbf{X}\mathbf{D}\mathbf{X} \succeq 0$ results in $\mathbf{X}_1\mathbf{D}_1\mathbf{X}_1 - \xi\mathbf{b}\mathbf{b}^\dagger \succeq 0$ and $\mathbf{b}^\dagger\mathbf{D}_1\mathbf{b} - \xi x^2 \geq 0$ [23]. From $\mathbf{b}^\dagger\mathbf{D}_1\mathbf{b} - \xi x^2 \geq 0$, we get $\mathbf{b}^\dagger\mathbf{D}_1\mathbf{b} \geq \xi x^2 > 0$ (hence $\mathbf{b} \neq \mathbf{0}$). It follows from $\mathbf{X}_1\mathbf{D}_1\mathbf{X}_1 - \xi\mathbf{b}\mathbf{b}^\dagger \succeq 0$ that $\mathbf{X}_1\mathbf{D}_1\mathbf{X}_1 \succeq 0$. With this, by using the assumption for k , we know \mathbf{X}_1 has rank one and can be expressed as $\mathbf{X}_1 = \mathbf{v}\mathbf{v}^\dagger$. Thus, $\mathbf{X}_1\mathbf{D}_1\mathbf{X}_1 = (\mathbf{v}^\dagger\mathbf{D}_1\mathbf{v})\mathbf{v}\mathbf{v}^\dagger$ and $\mathbf{v}^\dagger\mathbf{D}_1\mathbf{v} \geq 0$. We can prove that $\mathbf{v} = c_1\mathbf{b}$ for a certain scalar c_1 . To see why this is the case, let us rewrite $\mathbf{X}_1\mathbf{D}_1\mathbf{X}_1 - \xi\mathbf{b}\mathbf{b}^\dagger = (\mathbf{v}^\dagger\mathbf{D}_1\mathbf{v})\mathbf{v}\mathbf{v}^\dagger - \xi\mathbf{b}\mathbf{b}^\dagger$. Note that $\mathbf{v}^\dagger\mathbf{D}_1\mathbf{v} \geq 0$, hence according to Lemma 2, if $\mathbf{v} \neq c_1\mathbf{b}$, then $(\mathbf{v}^\dagger\mathbf{D}_1\mathbf{v})\mathbf{v}\mathbf{v}^\dagger - \xi\mathbf{b}\mathbf{b}^\dagger$ always has a negative eigenvalue. But this violates $\mathbf{X}_1\mathbf{D}_1\mathbf{X}_1 - \xi\mathbf{b}\mathbf{b}^\dagger \succeq 0$.

Thus, $\mathbf{v} = c_1 \mathbf{b}$. Next, we determine c_1 . Now we can write $\mathbf{X}_1 \mathbf{D}_1 \mathbf{X}_1 - \xi \mathbf{b} \mathbf{b}^\dagger = (\mathbf{b}^\dagger \mathbf{D}_1 \mathbf{b} |c_1|^4 - \xi) \mathbf{b} \mathbf{b}^\dagger$ and

$$\mathbf{X} \mathbf{D} \mathbf{X} = \begin{pmatrix} (\beta |c_1|^4 - \xi) \mathbf{b} \mathbf{b}^\dagger & (\beta |c_1|^2 - \xi x) \mathbf{b} \\ (\beta |c_1|^2 - \xi x) \mathbf{b}^\dagger & \beta - \xi x^2 \end{pmatrix} \quad (78)$$

where $\beta = \mathbf{b}^\dagger \mathbf{D}_1 \mathbf{b}$. Since $\mathbf{X} \mathbf{D} \mathbf{X} \succeq 0$, we have

$$\beta |c_1|^4 - \xi \geq 0, (\beta |c_1|^4 - \xi)(\beta - \xi x^2) \geq (\beta |c_1|^2 - \xi x)^2 \quad (79)$$

which results in $|c_1|^2 = 1/x$. With this, we can rewrite (76) as

$$\mathbf{X} = \begin{pmatrix} |c_1|^2 \mathbf{b} \mathbf{b}^\dagger & \mathbf{b} \\ \mathbf{b}^\dagger & x \end{pmatrix} = \frac{1}{x} \begin{pmatrix} \mathbf{b} \\ x \end{pmatrix} \begin{pmatrix} \mathbf{b} \\ x \end{pmatrix}^\dagger \quad (80)$$

which leads to \mathbf{X} has rank one. This completes the proof.

REFERENCES

- [1] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure communication over fading channels," *Security and Communication Networks*, vol. 54, no. 6, pp. 2470-2492, Jun. 2008.
- [2] P. K. Gopala, L. Lai and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Information Theory*, vol. 54, no. 10, pp. 4687-4698, Oct. 2008.
- [3] V. Aggarwal, L. Sankar, A. R. Calderbank, and H. V. Poor, "Secrecy capacity of a class of orthogonal relay eavesdropper channels," *EURASIP Journal on Wireless Communications and Networking, Special Issue on Wireless Physical Layer Security*, to appear.
- [4] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Physical layer security in broadcast networks," *Security and Communication Networks*, vol. 2, pp. 227-238, Wiley, May-Jun. 2009.
- [5] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz), "MIMO Gaussian broadcast channels with confidential messages," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Seoul, Korea, June-July 2009.
- [6] Y. Liang, H. V. Poor, and S. Shamai (Shitz), *Information Theoretic Security*, Now Publishers, Delft, The Netherlands, 2009.
- [7] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, pp. 1355-1387, Oct. 1975.
- [8] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Information Theory*, vol. 24, pp. 451-456, Jul. 1978.
- [9] A. Khisti and G. Wornell, "The MIMOME channel," in *Proceedings of the 45th Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, USA, September 2007.
- [10] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," in *IEEE International Symposium on Information Theory (ISIT)*, pp. 524-528, Toronto, ON, Canada, Jul. 2008.
- [11] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multi-antenna wire-tap channel," *IEEE Trans. Information Theory*, vol. 55, pp. 2547-2553, Jun. 2009.
- [12] R. Bustin, R. Liu, H. V. Poor, and S. Shamai (Shitz), "An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Seoul, Korea, June-July 2009.

- [13] S. Shafiee, N. Liu, and S. Ulukus, "Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel," *IEEE Trans. Information Theory*, submitted in 2007.
- [14] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel, " [online]. Available: <http://aps.arxiv.org/abs/0710.1920>, updated at Jul. 2009.
- [15] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound wire-tap channels, " in *Proceedings of the 45th Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, USA, September 2007.
- [16] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1165-1175, Jul. 2004.
- [17] M. Vu and A. Paulraj, "Optimal linear precoders for MIMO wireless correlated channels with nonzero mean in space-time coded systems," *IEEE Trans. Signal Processing*, vol. 54, no. 6, pp. 2318-2332, Jun. 2006.
- [18] Jiangyuan Li and Q. T. Zhang, "Transmitter optimization for correlated MISO fading channels with generic mean and covariance feedback," *IEEE Trans. Wireless Commun.*, vol. 7, no. 9, pp. 3312-3317, Sept. 2008.
- [19] V. Raghavan, V. V. Veeravalli and R. W. Heath, Jr., "Reduced rank signaling in spatially correlated MIMO channels, " in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Jun. 2007.
- [20] W. Dinkelbach, "On nonlinear fractional programming, " *Management Science*, vol. 13, no. 7, pp. 492-498, 1967.
- [21] S. Boyd and L. Vandenberghe, *Convex Optimization*, UK: Cambridge Univ. Press, 2004.
- [22] H. T. Davi and K. T. Thomson, *Linear Algebra and Linear Operators in Engineering*, Academic Press, 2000.
- [23] R. A. Horn and C. A. Johnson, *Matrix Analysis*, UK: Cambridge Univ. Press, 1990.

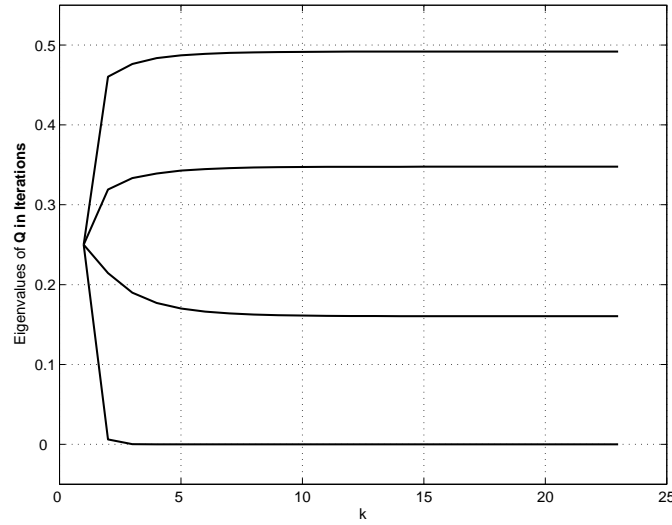


Fig. 1. Eigenvalues of \mathbf{Q}^k in iterations, positive definite case.

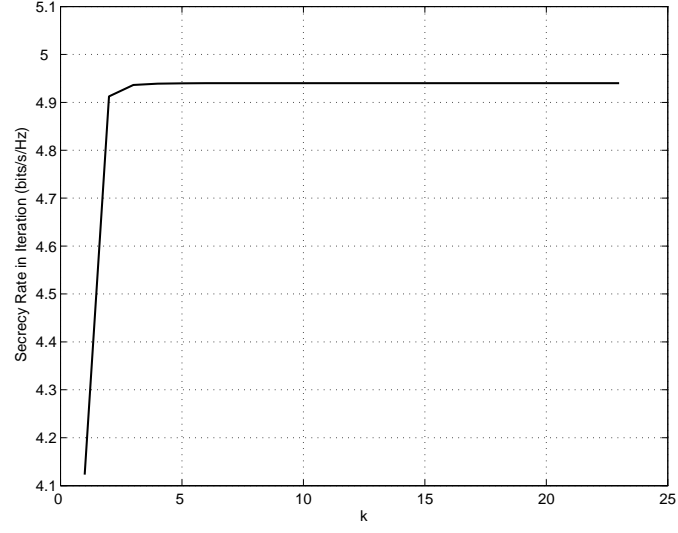


Fig. 2. Secrecy rate in iterations, positive definite case.

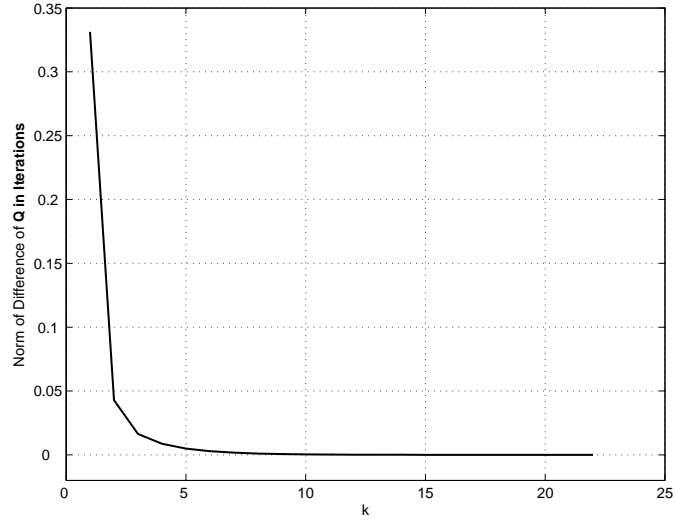


Fig. 3. $\|Q^{k+1} - Q^k\|$, positive definite case.

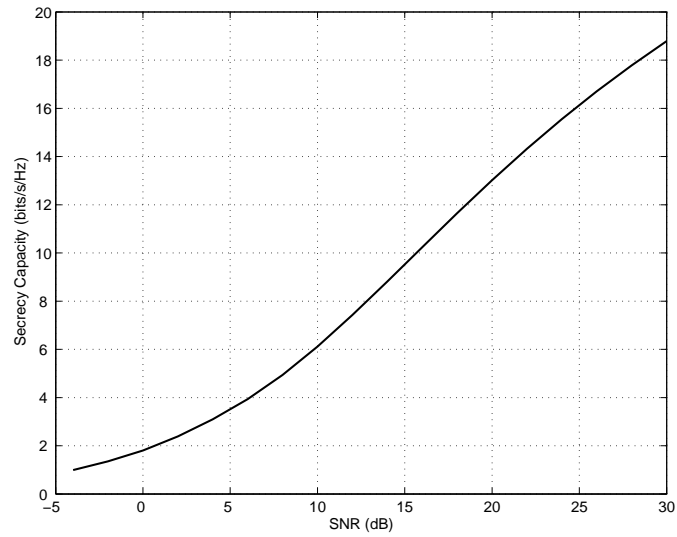


Fig. 4. Secrecy capacity Vs. SNR, positive definite case.

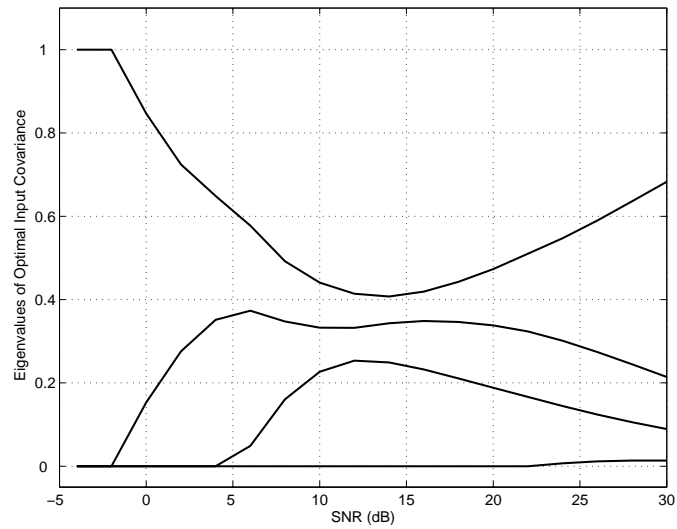


Fig. 5. Eigenvalues of optimal input covariance Vs. SNR, positive definite case.

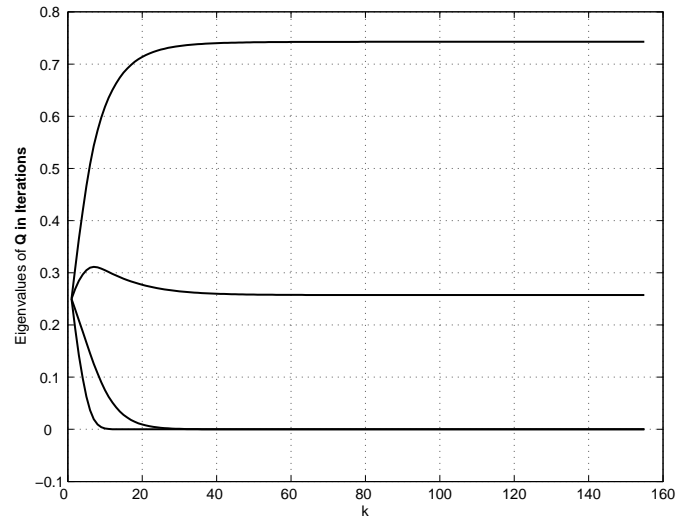


Fig. 6. Eigenvalues of Q^k in iterations, non positive definite case.

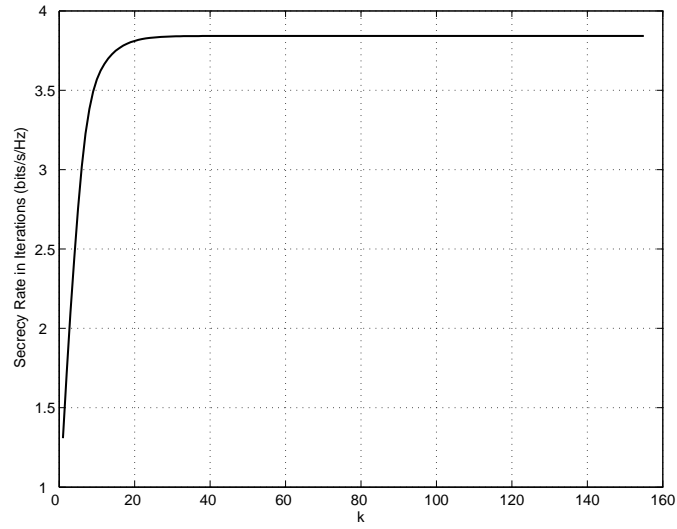


Fig. 7. Secrecy rate in iterations, non positive definite case.

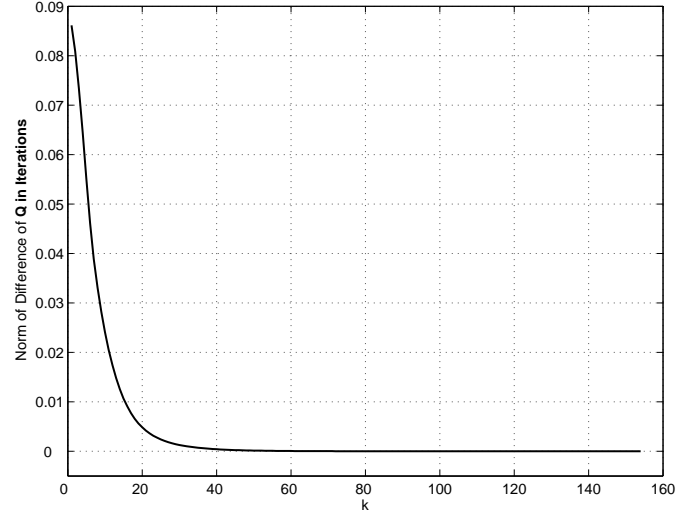


Fig. 8. $\|\mathbf{Q}^{k+1} - \mathbf{Q}^k\|$, non positive definite case.

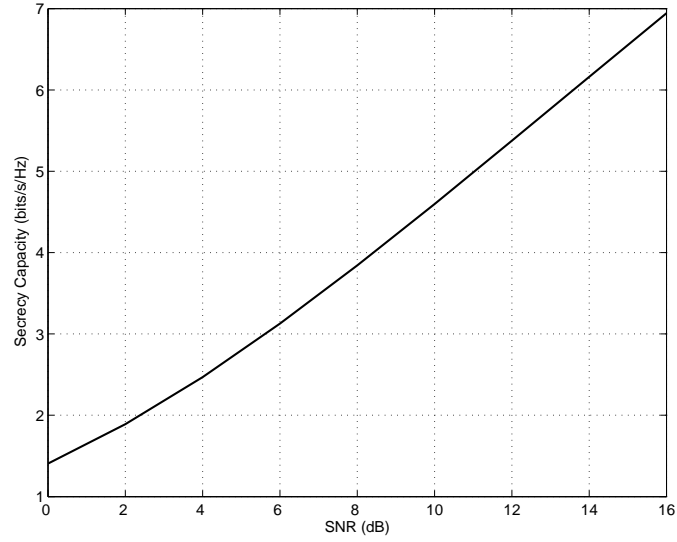


Fig. 9. Secrecy capacity Vs. SNR, non positive definite case.

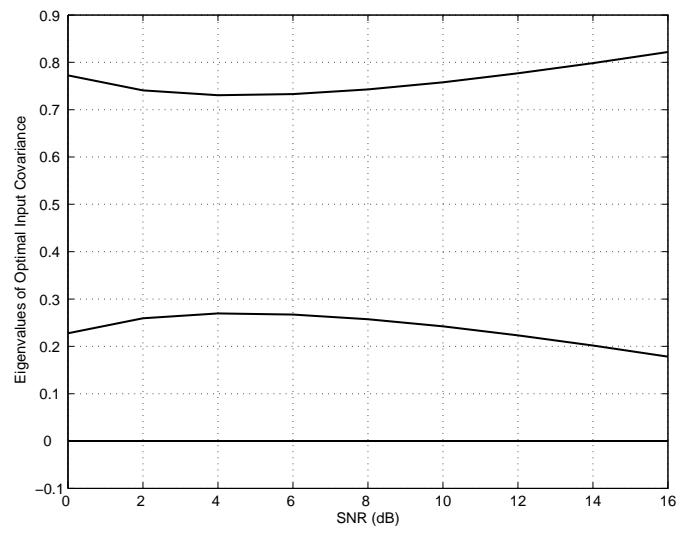


Fig. 10. Eigenvalues of optimal input covariance Vs. SNR, non positive definite case.